Audio recording do's and dont's

Coralie VINCENT
SFL, UMR 7023 - CNRS / Université Paris8

Sound parameters

Three characteristics:

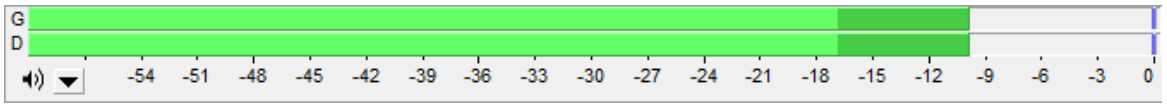
- (Fundamental) Frequency: F0 [Hz]
  (in the perception domain = pitch; ranges from 20 to 20000 Hz)

- Amplitude: measured via the sound pressure level [dB_{SPL}]
  (in the perception domain = loudness; ranges from 0 to 130 dB)

Other kinds of dB scale exist:

[dB(A)] (human ear): A-weighting takes into account the way our ears perceive and interpret sounds (especially at low and high frequencies)

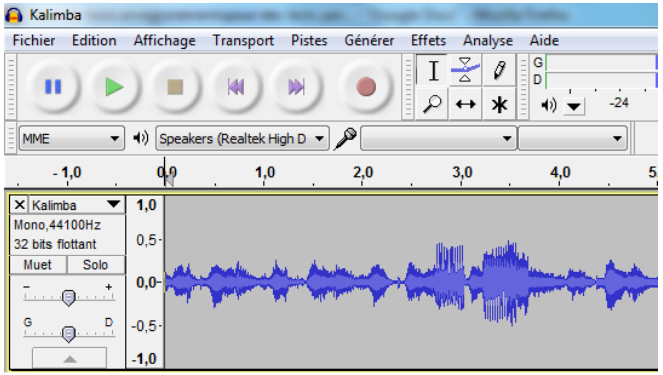
[dB_{FS}] (digital devices): used in digital devices (ranges from $-\infty$ to 0) Always < 0.



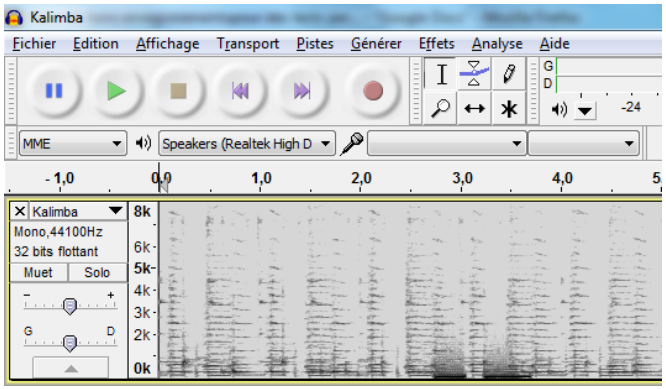
- Timbre is what you hear when the two first parameters are similar but there is still a difference (a piano and a flute sound, for example)

Sound representations

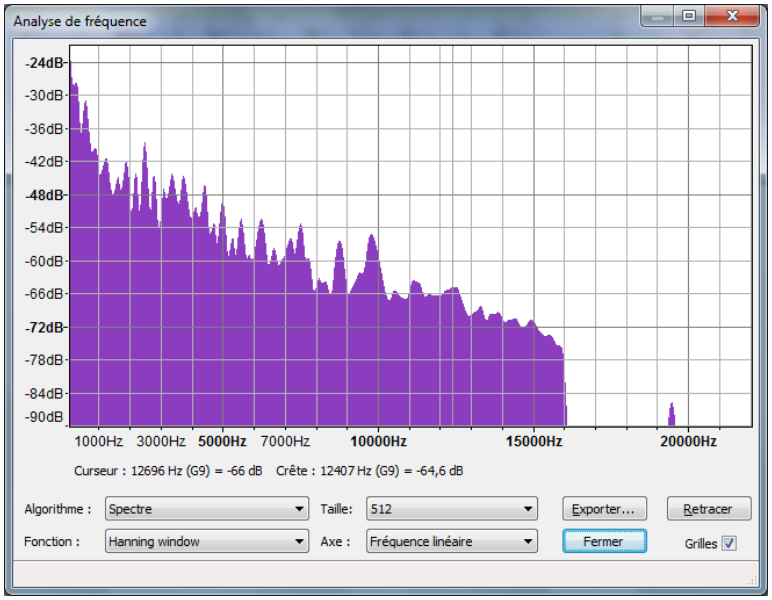
Waveform - amplitude vs. time.



Spectrogram - frequency vs. time, amplitude is more or less grey.



Spectral Slice - amplitude vs. frequency at a certain time.



Images from the free audio editor Audacity:

http://audacity.sourceforge.net/

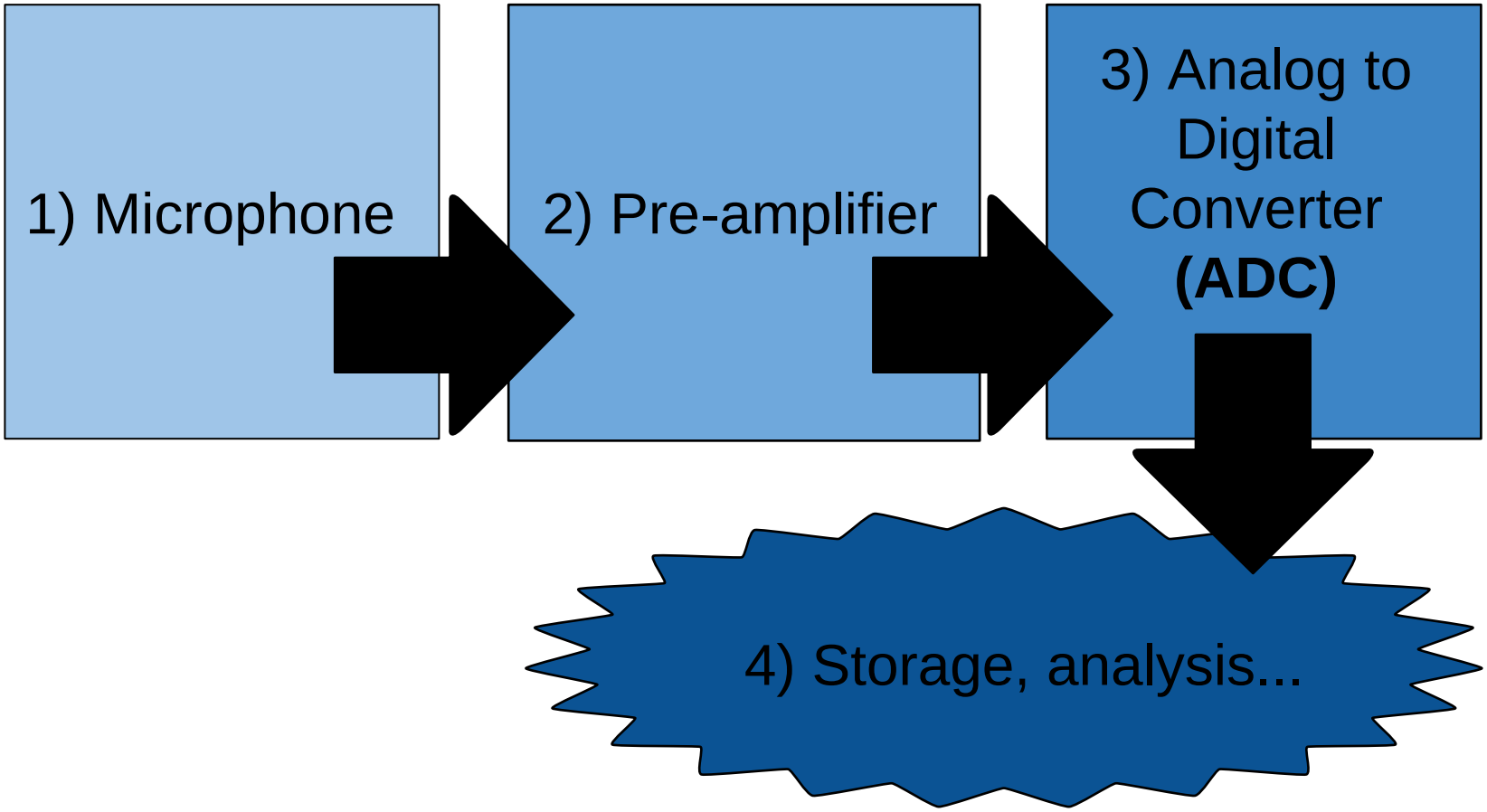
A few orders of magnitude for voice

- Frequency / timbre
  Minimum : 50 Hz (man) + beware of "creaky" voice
  Maximum : 8000 Hz (/s/).

- Dynamic range ( level depends on distance!!)
  Minimum : Whispering
  30 dBA at 30 cm (microphone on a stand)
  45 dBA at 5 cm (headworn microphone)
  Maximum : Scream, singing (opera)
  120 dB at 30 cm
  135 dB at 5 cm.

- Note: Distance / 2 → + 6 dB
  For example: if the distance from a microphone to a sound source is reduced from 30 cm to 15 cm → the level gains 6 dB.

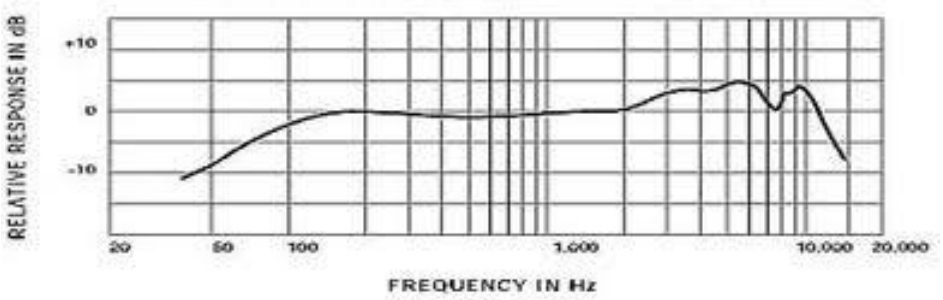
Acoustic signal path block diagram



1) Ideally... for a microphone[6]

- Flat frequency response from 20 to 20000 Hz.

- Actual response (relative response in dB vs. frequency in Hz):



1) Ideally... for a microphone,

- Dynamic range

  Considering extreme values for voice, we obtain:
  - Min. to match signal-to-noise ratio > to 15 dB
    - 15 dBA at 30 cm, 30 dBA at 5 cm
  - Max. to avoid distortion (soft or hard)
    - >120 dB at 30 cm, >135 dB at 5 cm

A perfect microphone for voice should be built to handle from 15 to 135 dB.

If the level is too high for the microphone, there are 2 types of distorsion:

- Soft distortion: microphone membrane (can be difficult to notice)
- Hard distortion: electronics, flattened peaks (audible).

1) Placing of a microphone Distance

Critical distance: distance where the direct and reverberant sound energies become equal.

Two cases for AIRS researchers:

1) For a "naturalistic" sound recording: place your microphones at the critical distance = good start.
2) For audio data which are going to undergo acoustical analysis (pitch, spectrogram, rhythm,...): the closer the better (to a certain point because of the proximity effect on cardioid microphones and to avoid distortion).

1) Placing of a microphone Direction

Voice directivity

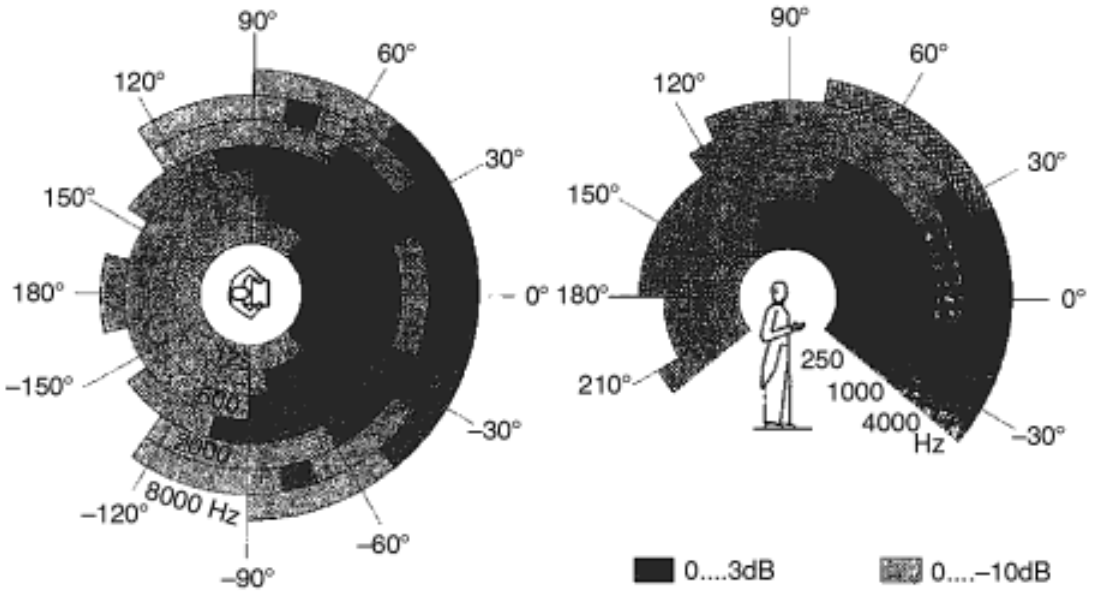


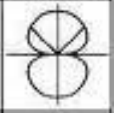
Fig. 4.33 Angular regions of principal radiation regions of the singing voice in octave bands (after Marshall and Meyer, 1985)

From "Acoustics and the Performance of Music" by J. Meyer (1978)

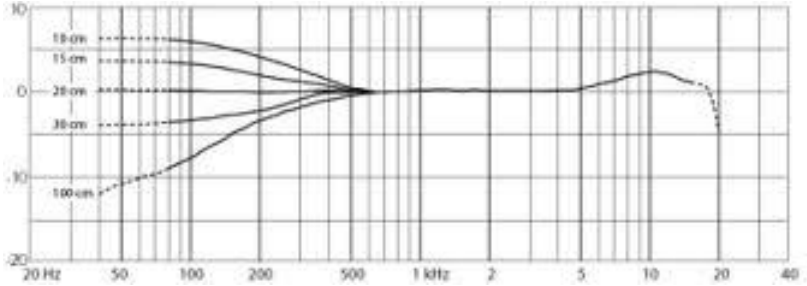
1) Placing of a microphone Orientation & proximity effect

- Directivity diagram

Omnidirectional vs. cardioid

CHARACTERISTIC	OMNI-DIRECTIONAL	CARDIOID	SUPER-CARDIOID	HYPER-CARDIOID	BI-DIRECTIONAL
POLAR RESPONSE PATTERN					
COVERAGE ANGLE	360°	131°	115°	105°	90°
ANGLE OF MAXIMUM REJECTION (NULL ANGLE)	— °	180°	126°	110°	90°
AMBIENT SOUND SENSITIVITY (RELATIVE TO OMNI)	100%	33%	27%	25%	33%

- Proximity effect (only for a cardioid):



1) Microphone types (dynamic vs. electrostatic)

Dynamic

- Coil & magnet
- No need for power supply
- Frequency response < electret & condenser
- Not sensitive

Dynamic:

less expensive than Electret

Electret / condenser (also called electrostatic)

- Membrane : condenser
- Frequency response flatter
- Need for power supply (battery or "phantom",48 V)
- Sensitive

Electret:

less expensive than condenser

Advice: use an electrostatic microphone unless you are doing field research.
Dynamic microphones are useful if you need a robust microphone or if you are facing severe weather conditions (very humid, very cold,…).

Connectors (mind the buzz) 12

Connectors:

XLR

Cinch (RCA) red = right

TRS jacks

3,5 mm

(1/8 inch)

6,35 mm

(1/4 inch)

What is a hum sound like?

Ex. 1 Ex. 2

Avoid power supply cables running along audio cables on a long distance

2) Preamplifier

2 roles :

- Impedance matching (check that input impedance is > 10 X microphone impedance > load impedance indicated in the microphone notice)
- Amplification by a factor of 100 to 1000 (from ~ 1-10 mV to ~ 1 V). Preamp boosts the low-level signal from the microphone (sensitive to noise and interferences) to an intermediate level used for digitization.

Dynamic range & frequency response curve must include those of the microphone.

The preamplifier is often integrated in the same module as the converter (portable recorders, soundcards,...).

3) ADC and recorder (mind the buzz again)

A/D Converter: 2 parameters:

- Sampling frequency:

  Nyquist-Shannon Theorem : $F_S \geq 2 \times F_N$ (Nyquist freq.)

  Normalized frequencies: 44,1 kHz (CD) & **48 kHz**
  (48 kHz recommended for files uploaded in the Digital Library)

- Bit-depth:

  16 bits (CD): theoretical dynamic range of 96 dB

  24 bits: theoretical dynamic range of 144 dB (1,5X16 bits)

  (24 bits recommended for files uploaded in the Digital Library)

Portable recorder: if possible, run on batteries to avoid 60 Hz hum (50 Hz in Europe).

4) Storage (DL) - File size

Uncompressed audio ("PCM" wav, aiff):

48 ([kHz]) * 24 ([bits]) * 2 (stereo) = 2304 kbit/s

Per hour: 2304 * 3600/8 = 1036800 ko ~ **1 Go**

(.wav file unavailable in .pdf - 320 kbps mp3 file instead)

Comparison with compressed audio (mp3):

128 kbit/s

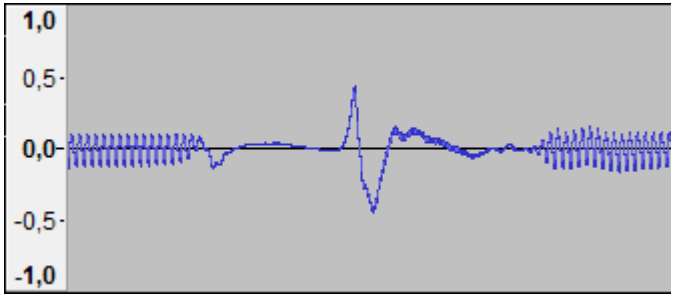
Per hour: 128 * 3600/8 = 57600 ko ~ **56,25 Mo**

What is lost through lossy audio compression? +40dB:

NEVER record in a compressed format (aac, mp3...).
Beware of MD & Hi-MD too.

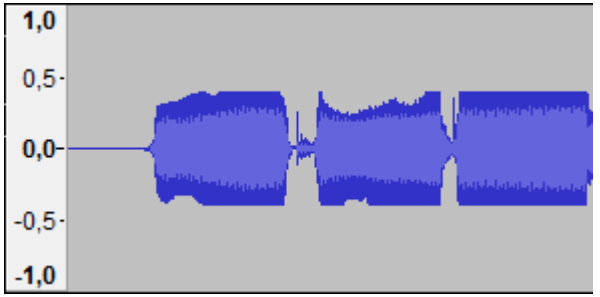
In practice - Waveforms

Pop

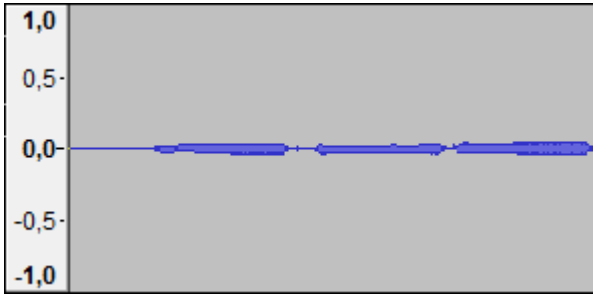


“mon ami Pierrot”

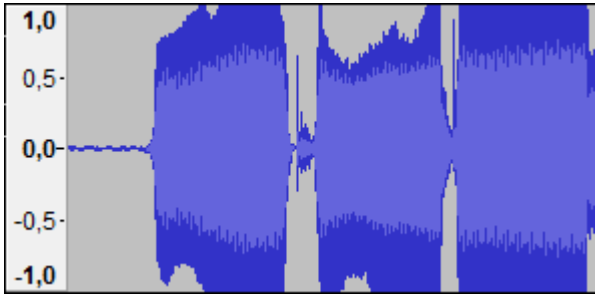
Clipping (soft / hard)



Undermodulation



Clipping (hard)



In practice

- Check that the equipment is adapted (connectors, impedance...).
- Avoid putting the recorder (which might have vibrating elements) on the table where the microphone stand is. A suspension for the microphone also limits these vibrations.
- If you hear pops, use a *pop filter* → or put the microphone off to the side.

- Don’t put your headphones on until everything is plugged…
- **Speech**: while adjusting the various gain levels, try to settle the average voice level at ~ - 20 dB_{FS} on the digital recorder.
- **Singing**: ask the performers to sing the climax of their piece. Make sure there is still “room” (~ 6 dB) above this climax.

In practice

- Adapt your sound recording setup to your environment (equipment choice) & be conscious of each device's limit.
- Turn off cellphones and noise sources (air conditioning, wind,...).
- Use a sentence with important levels to setup gains.
  ("Paul, please pause for proper applause")
- At the beginning of a session, record an excerpt and listen to it before recording for a long time.
- Don't forget to press the "REC" button! 
- If you use Praat to view your recorded sounds, beware of the autoscale function which may make an undermodulated sound look beautiful.

Bibliography

Consult microphone manufacturers' documentations & websites (AKG, Audio Technica, Beyerdynamic, Brüel & Kjaer, DPA, Sennheiser,…)

Baken R. J., & Orlikoff, R. F. (2000). *Clinical measurement of speech and voice* (2e éd.). San Diego, CA : Singular.

Shannon, C. E. (1949). Communication in the presence of noise. *Proceedings of the IRE, 37,* 10-21.

Švec, J. G., Granqvist, S. (2010). **Guidelines for Selecting Microphones for Human Voice Production Research.** *American Journal of Speech-Language Pathology, 19,* 356-368. This paper is HIGHLY RECOMMENDED!!!

Ternström, S., & Granqvist, S. (2010). Personal computers in the voice laboratory: Part two - audio devices. *Logopedics Phoniatrics Vocology, 35,* 98-102.

Acknowledgements and useful websites

I would like to thank my friend and colleague Claire Pillot-Loiseau who accepted to “lend me” her voice. Theresa Leonard’s advices were also very helpful.

Interesting websites:

http://www.auditory.org/ (especially the page on portable audio recorders: http://www.auditory.org/recorders.html)

http://bartus.org/akustyk/

http://vocalmicrophonepro.com/

http://en.wikipedia.org/wiki/Microphone_practice

http://en.wikipedia.org/wiki/Audio_file_format

Appendix: Mono & 2-Track vs. Stereo

Mono

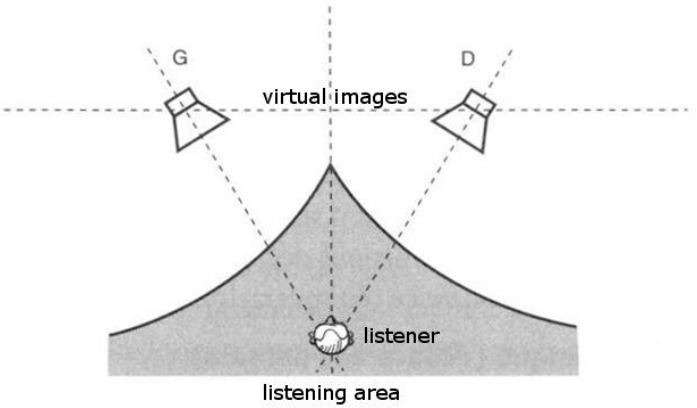
* One microphone recorded on one track.
* If played on headphones, both earflaps play exactly the same sound

2-Track

* Two distinct microphones recorded on two tracks
* If played on headphones, earflaps play distinct sounds

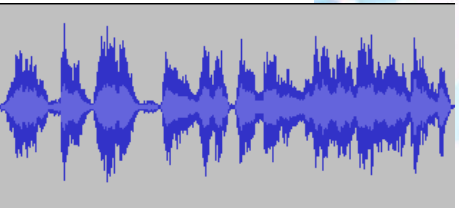
Stereo

* Two microphones (couple) with correlated signals recorded on two tracks
* Renders space
* Needs appropriate sound system to be reproduced (2-loudspeaker setup)



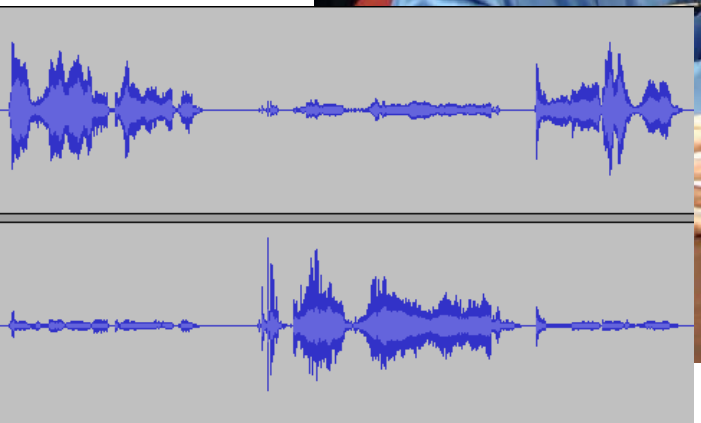
Mono Setup

Waveform:



2-Track Setup

Waveform:



Stereo Setup

Waveform: