

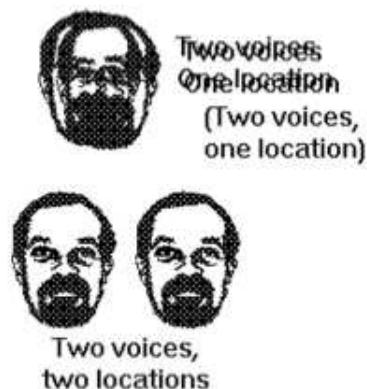
Music 175: Time and Space

Tamara Smyth, trsmyth@ucsd.edu
Department of Music,
University of California, San Diego (UCSD)

April 16, 2019

The “Cocktail Party Effect”

- **Cocktail Party Effect:** ability to follow speech from one speaker in the presence of chatter from many others.
- What allows us to distinguish speech from one speaker when two or more are speaking simultaneously?
- **Location:** tuning in on a particular speaker is easier when speakers are in different locations.



- **twoVoicesLocation.pd:** Two different messages voiced by the same person are played simultaneously:
 1. with both through both channels,
 2. each through a different channel.
 - In which is the message easier to decipher?

Binaural Masking

- Though signals arrive at the ears with different phases, the ear has the ability to adjust the travel times in neural pathways from the two ears so signals **add constructively**.

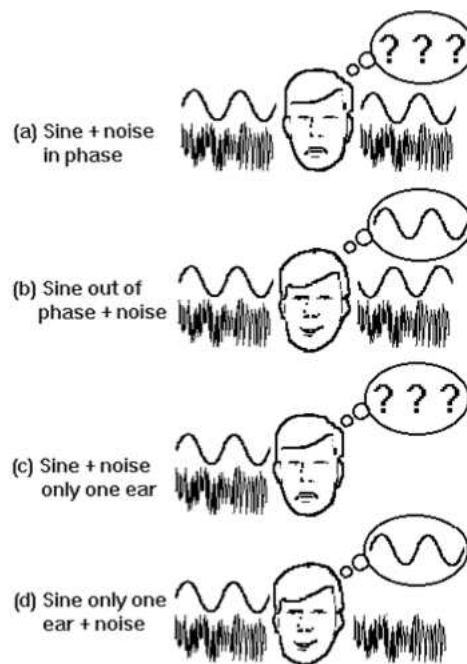


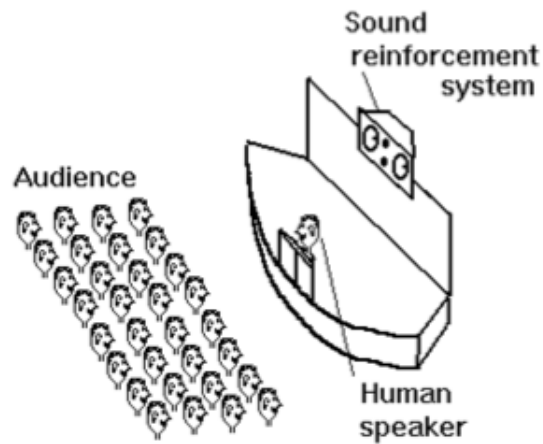
Figure 1: If the sine is different in the two ears and the noise is not, it's easier to hear.

- Experiment: [binauralMasking.pd](#)
- This ability of the auditory system to shift the relative times at which the signals of the two ears are added helps us “tune in” (essential to cocktail party effect).

The Precedence Effect

- Two identical sounds played in close succession will be heard as a single **fused sound**.
 - $1 \text{ ms} < \text{lag} < 5 \text{ ms}$: effect heard with clicks;
 - $\text{lag} < 40 \text{ ms}$: heard with more complex sound;
 - $\text{lag} > 40 \text{ ms}$: second sound is heard as an echo
- **Precedence Effect**: the perceived location of successive sounds heard as fused but coming from *different locations*, is dominated by the location of the **sound that first reaches the ears** (first-arriving wavefront).
 - in a room with reflecting surfaces, a combination of many reflected sounds may reach our ear with greater intensity than direct sound;
 - the tendency is to hear the sound as coming from the direction from which it first reaches the ears.
- Experiment: Play voice from two speakers: from where do you hear the voice?
- [precedence.pd](#)

- It is common to place an amplification system *behind* the speaker so that the sound is perceived as coming from the speaker.



Reverberation

- Reverberation is produced naturally by the reflection of sounds off surfaces.

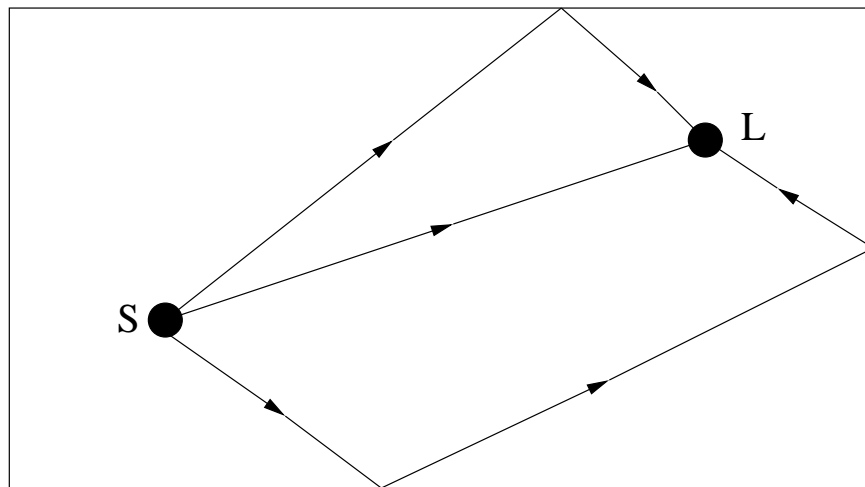


Figure 2: Example reflection paths occurring between source (S) and listener (L).

- It's effect on the overall sound that reaches the listener depends on the room or environment in which the sound is played.
- **Impulse response of a cave**

Reflections

- There are several paths the sound emanating from the source can take before reaching the listener.
- The closest path is the one taken by the *direct sound*.
- Delayed images reaching the listener lengthen the time the listener hears the sound.
- Recall, the amplitude of the sound decreases at a rate inversely proportional to the distance traveled.
 - sound is not only delayed, but it also decays.
 - reverberation tends to have an (mostly exponentially) decaying amplitude envelope.
- Four physical measurements that effect the character of reverberation (and thus character of the space):
 1. reverb time (RT or T60);
 2. frequency dependence of RT;
 3. time delay between the arrival of the direct sound and the first reflection;
 4. rate at which the echo density builds.

Reverb time or T60

- The reverb time or T60 is
 - how long a listener will hear a sound,
 - the time required for a sound to decay by 1/1000 (or level by -60 dB).
- T60 depends on:
 - **Volume**: large volume rooms tend to have longer T60s.
 - **Surface area**: with constant volume, T60 will decrease with an increase in surface area available for reflections (and thus absorptions).
 - **Nature of surface area**:
 - * **absorptivity**: soft porous surfaces (curtains, carpet, upholstered chairs) absorb more acoustic energy than hard, solid, nonporous surfaces.
 - * **roughness**: if the surface is not perfectly flat, part of the sound is reflected and part is dispersed in (many) other directions.
- T60 is also dependent on the amplitude of the original sound and the presence of other sounds.
- Listen to **RT demo**.

Frequency Dependence of Reverb Time.

- Reverb time is not uniform over audible frequencies:
 - in a well designed concert hall, the low frequencies are the **last** to fade.
 - absorptive materials tend to reflect better at low frequencies
 - hard, nonporous reflectors (such as marble) reflect sounds of all frequencies with nearly equal efficiency.
- With small solid objects, the **efficiency** and the **direction** of reflections are both dependent of frequency (wavelength):
 - this causes frequency-dependent dispersion and an alteration of the waveform of a sound.

Delay Between Direct Sound and First Reflection

- A long delay (> 50 ms) can result in distinct echoes.
- A short delay (< 5 ms) contributes to the listener's perception that the space is small.
- A delay between 10 and 20 ms is found in most good halls.

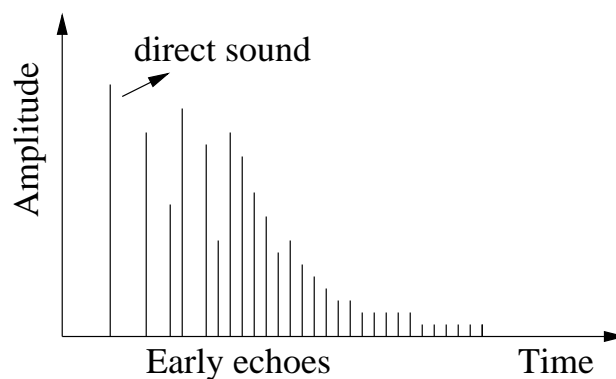


Figure 3: An example impulse response showing direction sound, early reflections, and exponential decay of the sound.

Rate at which Echo Density Builds

- After the initial reflection, the rate at which echoes reach the listener increases rapidly.
- A listener can distinguish differences in **echo density** up to a density of 1 echo/ms:
 - the amount of time required to reach this threshold influences the character of the reverberation (100 ms in a good situation).
 - this time is roughly proportional to the square root of the volume of a room (small spaces are characterized by rapid buildup of echo density).

Interaural Coherence

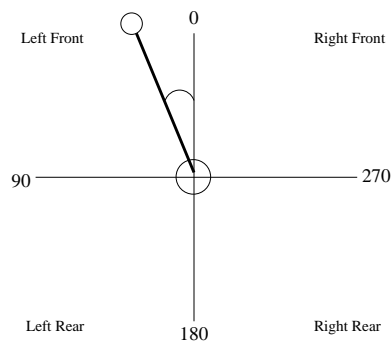
- Each acoustic space has its own variety of reflectors causing the reverberant sound arriving at the listener to be different from each direction.
- *Interaural coherence* is a measurement that indicates similarity between the reverberation received by each of the two ears.
- A low interaural coherence generally results in a more pleasing sound and a greater feeling of “immersion”.
- Low interaural coherence can be implemented by giving each channel its own reverberator with slightly different parameters.

Auditory Localization

- Both reverberation and locality add dimension to sound.
- **Auditory localization** is the human perception of the placement of a sound source.
- A listener receives cues indicating a sound's placement in an acoustic space (e.g. a room or concert hall).
- Sound from a loudspeaker sounds like it's coming from a loudspeaker:
 - *imagery* in an audio production can create the difference between “a violinist in a room” and a “loudspeaker reproduction”;
- The **location** of a sound source is typically defined by its
 - **direction**;
 - **distance**.

Direction Expressed as Angles

- The **direction** is usually expressed in terms of angles.
- **Azimuth angle ϕ**
 - measured in the horizontal plane passing through the center of the listener's head.
 - determines the position of the source in the four quadrants surrounding the listener's head: $\phi = 0^\circ$ is situated directly in front of the listener; $\phi = 180^\circ$ is situated directly behind.



- **Angle of elevation θ**
 - measured in a vertical plane bisecting the listener: $\theta = -90^\circ$ is situated directly below the listener; $\theta = +90^\circ$ is situated directly above.

Primary Localization Cues

- The **direction** is usually determined by **time** and **intensity** differences, as received by the two ears.
- The pinnae filters the sound in a way that is **directionally** dependent:
 - particularly useful in determining if a sound comes from above, below, in front, or behind.

Interaural time difference (ITD)

- The ITD is the delay that a listener perceives between the time that a sound reaches one ear and the time that it reaches the other.
- The ITD cues give information regarding the angular direction of a source source:
 - if the source is directly in front or behind the listener, the sound will reach both ears at the same time and the ITD will be zero.
 - a typical listener can resolve the location of a sounds in front to about 2° and from behind to about 10° .

Interaural intensity difference (IID)

- When the sound source is not centered, the listener's head partially “shadows” the ear opposite to the source, diminishing the intensity of the sound in that ear
 - IID is frequency dependent and increases with frequency

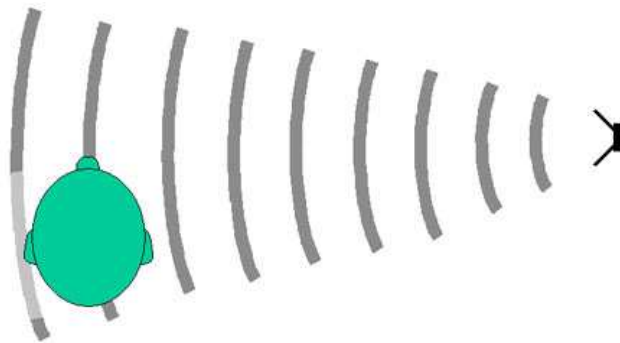


Figure 4: Acoustic shadow contributes to IID.

ITD and IID Cues

- The degree to which these cues are effective is dependent on the **frequency** content of the sound.
- Both ITD and IID are ineffective at low frequencies (below 270 Hz), and thus the direction of such sounds is more difficult to determine.

ITD Cues

- less precise behind a listener because the change in the ITD per degree in location change is smaller.
- most effective between 270 and 500 Hz
- little contribution above 1400 Hz

IID Cues

- less sensitive to sound sources behind the head.
- very small for frequencies below 500 Hz
- contribute more at higher frequencies
- dominate at (and above) about 1400 Hz

Other Localization Factors

- IID and ITD are insufficient localization cues.
- Pinnae filtering is most pronounced above 4 kHz.
- **Reverberation** can also provide a localization cue but works best on impulsive sounds.
- Longer tones are more difficult because listeners estimate distance almost entirely during the attack portion of the sound (where ITDs are most effective).
- Reflections off the torso, shoulder also serve a cues.
- Mistakes in localizing the sound occur mostly at low frequencies.

Head Related Transfer Functions

- To increase fidelity of localization, researchers have measured *head-related transfer functions* (HRTFs) over a wide range of incidence angles.
- HRTFS express the frequency response imparted to a sound by the pinnae, head, shoulders, torso etc., for a particular angle.
- HRTFs are also dependent on distance, though little variation is observed for source locations more than 2m away (making this the common distance for HTRTF measurements).

Simulation of Directional Cues

- Directional cues may be provided by
 1. using several loudspeakers;
 2. creating an *illusion*.
 - a listener positioned equidistant from two loudspeakers L and R receives equal signal in each ear;
 - the *illusion* is of a source centered at I1.

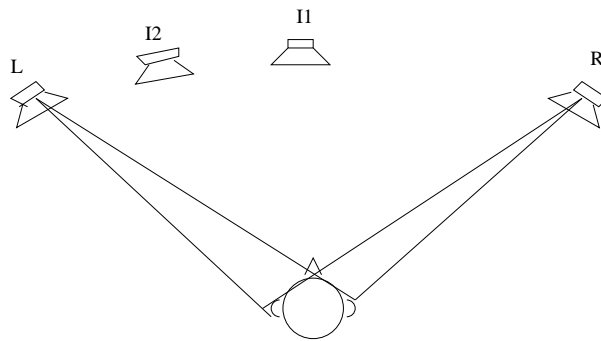


Figure 5: Listener positioned equidistant from two loudspeakers.

- delaying the signal applied to loudspeaker R (equivalent to moving R farther to the right) causes the sound to reach the left ear first;
 - the *illusion* is of a source location shifted to I2.
- Simulating ITD cues requires very contrived conditions (e.g. headphones or a fixed position).

Cross Talk

- The sound that reaches an ear from the opposite loudspeaker is called *cross-talk*.

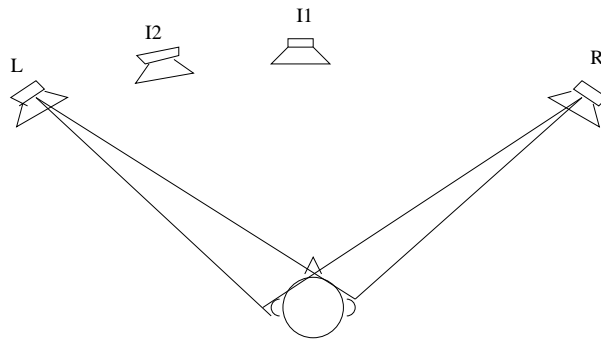


Figure 6: Listener positioned equidistant from two loudspeakers.

- Cross talk effectively limits the placement of auditory images to the area between the speakers.
- It is possible to compensate for cross-talk using filters if the listener is accurately positioned (e.g. wearing headphones).

Stereo Method for Simulating Localization

- The direction of the sound may also be simulated by changing the relative intensities at the speakers.

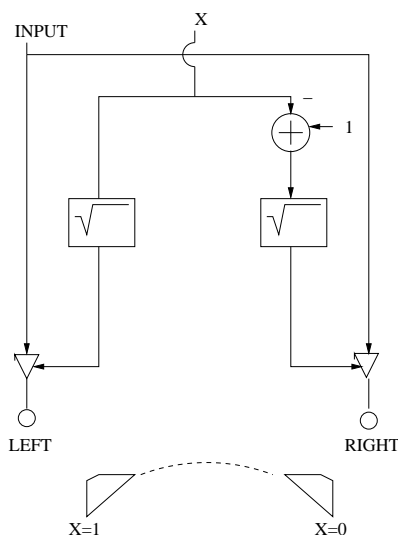


Figure 7: Stereo method for simulating localization cues.

- The location may be specified by parameter x .
- The **power** in the signal is allocated according to the value of x :
 - the power in the left speaker is given by x
 - the power in the right speaker is given by $1 - x$.
- Because x represents power, its square root is applied to the amplitude.

The principle cues for judging distance

- **Intensity of the sound**

- amplitude diminishes inversely with distance
- depends on listener's familiarity with the sound

- **The ratio of reverberated to direct sound**

- when the source is close to the listener, ratio of reverb to direct (R/D) is low.
- reflected energy increases with an increase in distance.
- at very large distances, an *audio horizon* is reached, and distance cannot be discerned.

- **Amount of high-frequency energy in the sound**

- attenuation of a sound wave propagating through the atmosphere is greater at high frequencies.
- at long distances there is perceivable absence of high-frequency components.

Global vs. Local Reverberation

- *Global reverberation* returns equally from all directions around the listener
- *Local reverberation* comes from the same direction as the direct signal and derives from reflectors relatively near the source.
- When the sound is located close to the listener, most of the reverberation is **global**.
- When the sound is located at a greater distance from the listener, most of the reverberation is **local**.