Auditory Scene Analysis

Week 9

Otherwise known as

Auditory Grouping

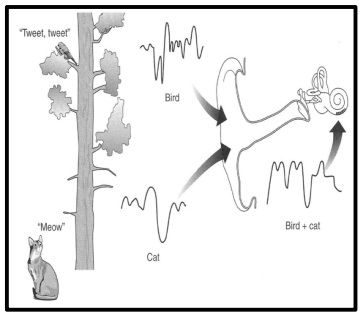
Auditory Streaming

Sound source segregation

Assigning acoustic/auditory features to distinct objects or sources of sound

The auditory scene

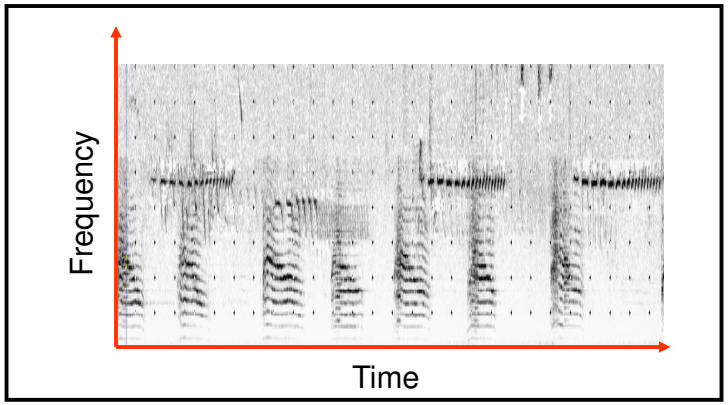
- The auditory system needs to make sense of the superposition of component sounds – the ***auditory scene***.
- It needs to ***segregate*** the components of the sound that come from different sound sources.
- It needs to ***group*** the components of the sound that come from the same sound source.

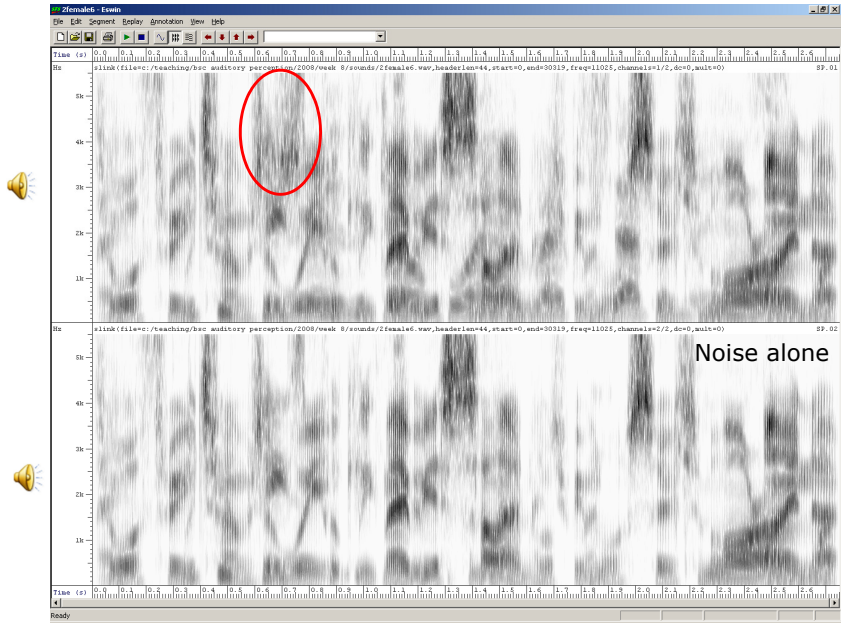


- The percept of a group of sequential and/or simultaneous sounds as a coherent whole appearing to come from a single sound source is known as a ***stream*** or ***auditory stream***.

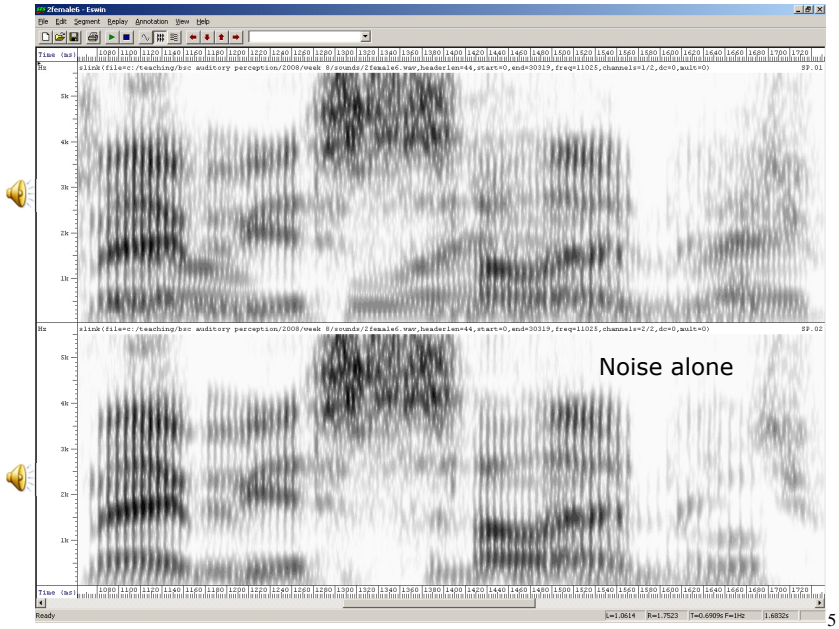
The auditory scene

Whole auditory scene = Cat (1st stream) + Birds (2nd stream)



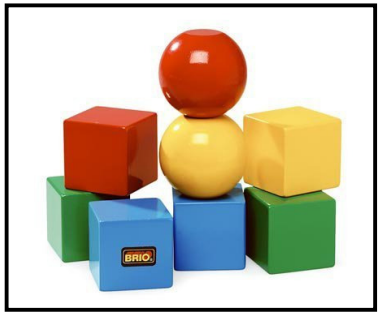


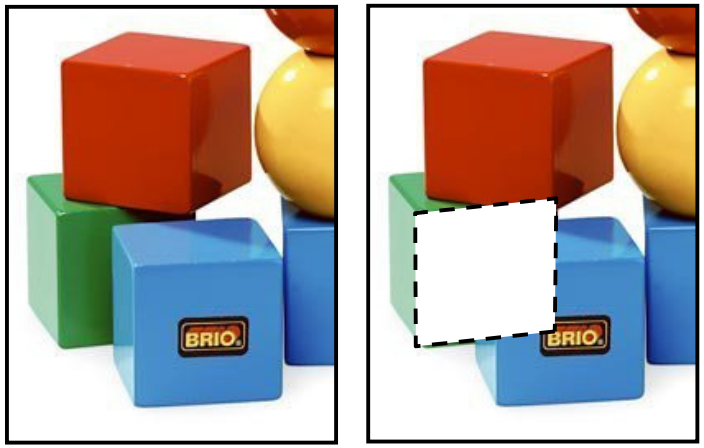
'show' starts at t≈0.65 ms



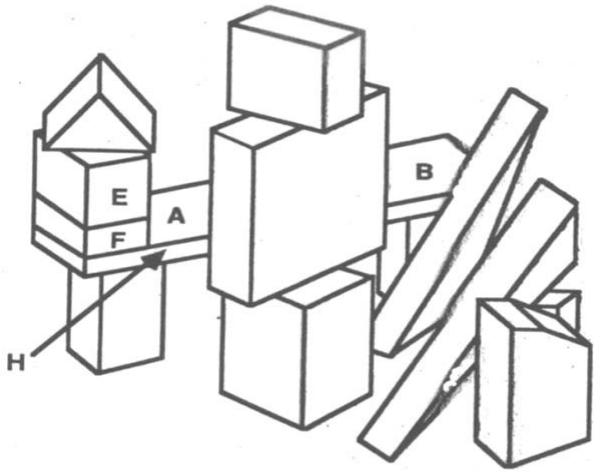
Visual scene analysis

- The principles of auditory scene analysis are the same as for visual scenes.
- How do we know what parts of the visual scene correspond a single object?
- How do we know what parts of a visual scene correspond to different objects?





The visual analogue: Assigning visual features to distinct objects



Visual scene analysis

- The principles for visual scene analysis were proposed by ***Gestalt*** psychologists in the early 20th century.
- They proposed a set of ***Gestalt grouping rules*** that describe which elements in an image belong together to form an object.
- These principles function so that our perceptual world is organized into the simplest pattern consistent with sensory information and our experience.
- Application of these principles *together* generally results in a grouping of the parts of an image that come from the same object and segregating those that don't.

Visual examples of Gestalt principles

Law of **Prägnanz**

Reality is organized or reduced to the simplest form possible.

For example, we see the image above as a series of circles rather than as many much more complicated shapes.

Law of Similarity:

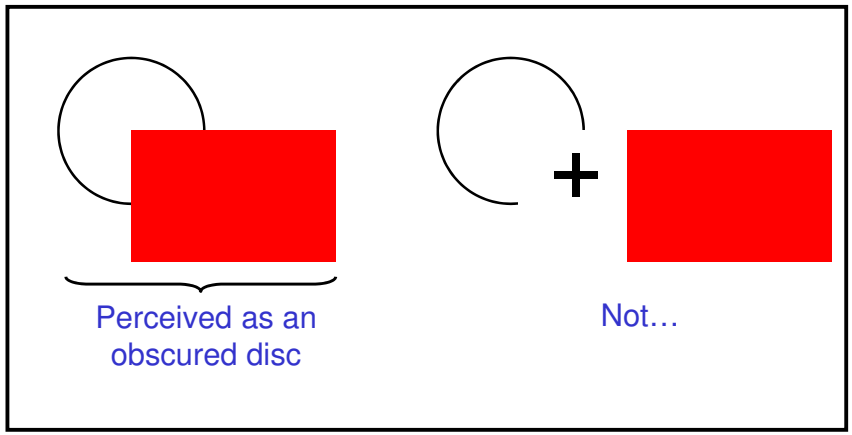
Items that are similar tend to be grouped together.

In the image above, most people see vertical columns of circles and squares.

http://psychology.about.com/od/sensationandperception/ss/gestaltlaws.htm

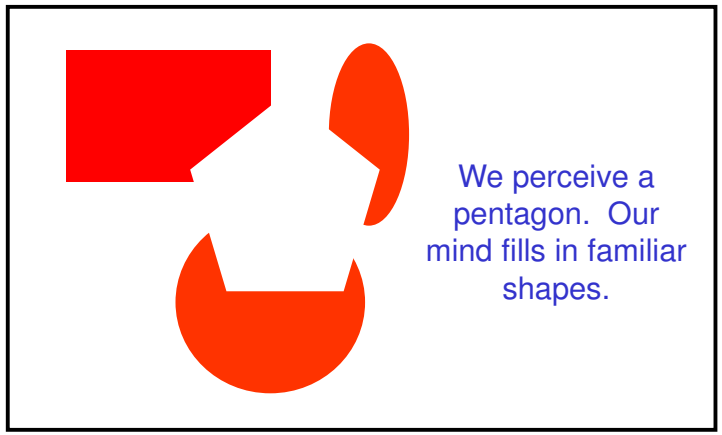
Visual completion by closure

- We tend to see completed or closed figures from contours, even when they are incomplete or open.



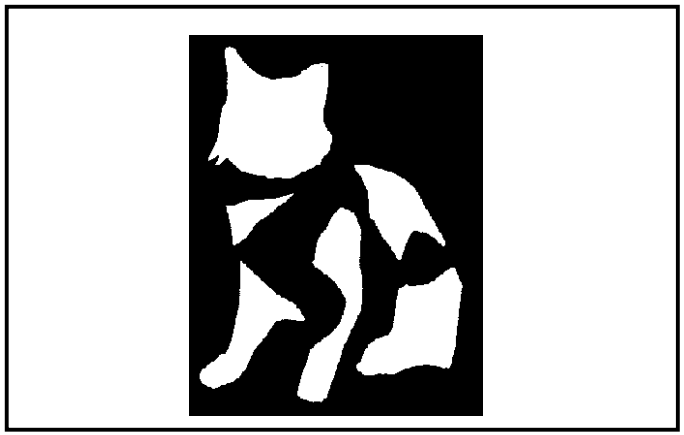
Visual completion by closure

- Perception is a constructive process – it's an interaction of stored knowledge and incoming sensory information.

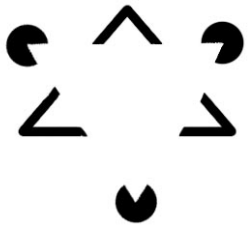


Visual completion by closure

- A whole cat and not disconnected shapes



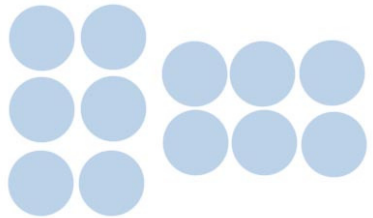
Visual examples of Gestalt principles



Law of Closure:

Objects grouped together are seen as a whole.

We tend to ignore gaps and complete contour lines. In the image above, there are no triangles or circles, but our minds fill in the missing information to create familiar shapes and images.



Law of Proximity:

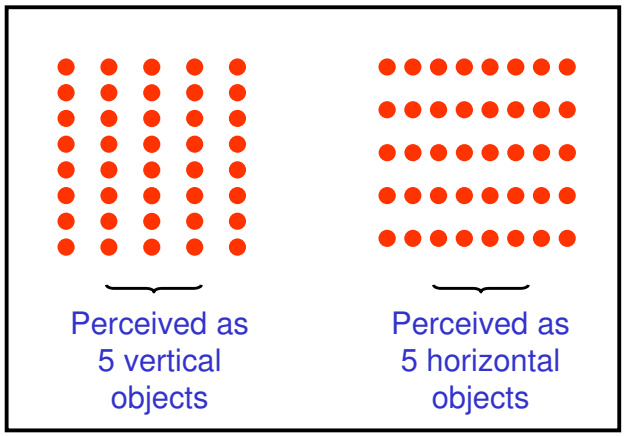
Objects near each other tend to be grouped together.

The circles on the left appear to be grouped in vertical columns, while those on the right appear to be grouped in horizontal rows.

http://psychology.about.com/od/sensationandperception/ss/gestaltlaws.htm

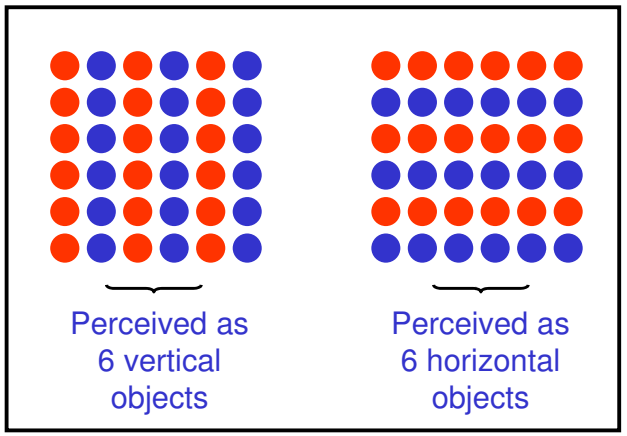
Visual grouping by proximity

- Things close together are perceived as one group.



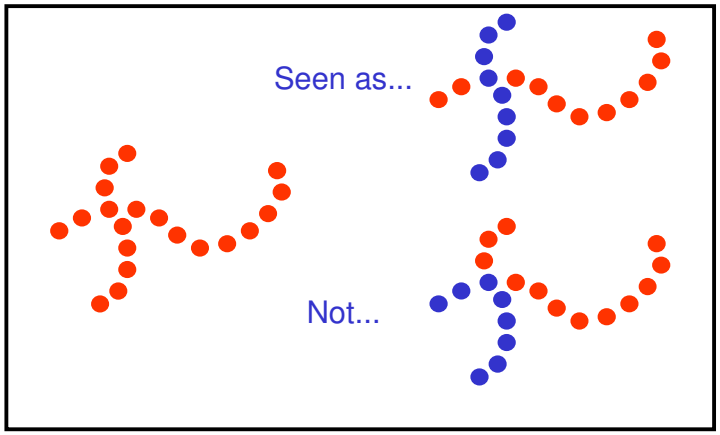
Visual grouping by similarity

- Similar things are perceived as one group.



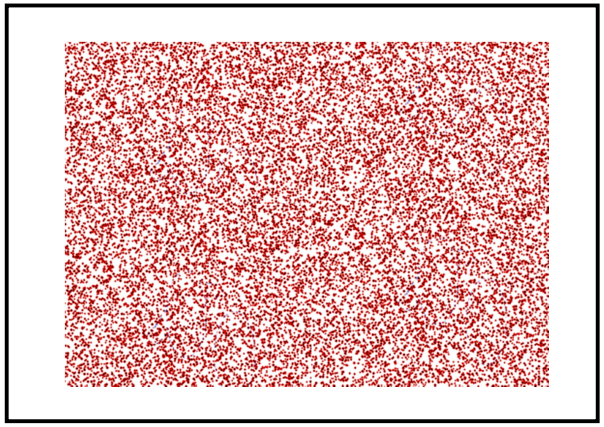
Visual grouping by continuity

- Lines are seen as following the smoothest path.



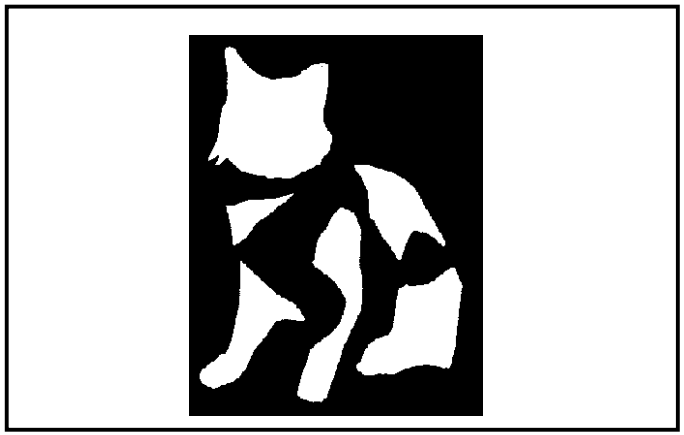
Visual grouping by common fate

- We tend to group things that are moving in the same direction and with the same velocity.



Separation in figure and ground

- We tend to organize our perceptions by distinguishing between a figure and a ground.
- Attention is generally focussed on the figure.



Gestalt principles

- Proximity
- Similarity
- Continuity
- Closure
- Common fate
- Disjoint allocation
  - An element of a visual scene must belong to a single object.
- Figure/ground

Auditory scene analysis

Sequential and simultaneous grouping

- Some cues enable sequential grouping of segments (at a temporal or melodic level) into separate streams **(across time)**
  - Proximity (pitch, time, location)
  - Similarity (timbre, loudness)
  - Continuity of pitch, loudness, location
- Some cues enable simultaneous grouping of segments (at a spectral or harmonic level) into separate streams **(across frequency)**
  - Harmonicity
  - Common fate: coherent changes in frequency, loudness, spectral envelope

Auditory streaming

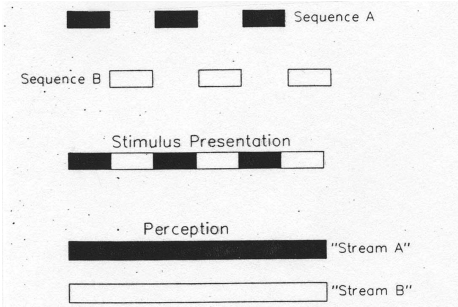
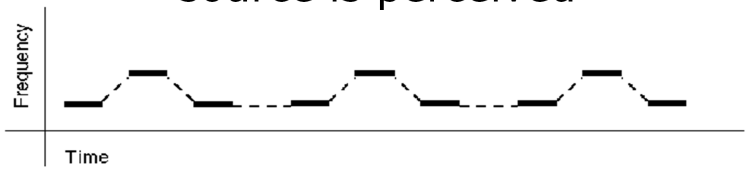


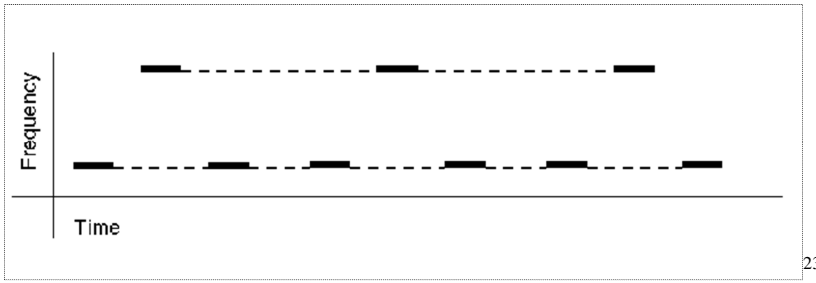
FIGURE 15.5 A schematic diagram indicating the type of procedure used in many streaming experiments. Two alternating sounds (e.g. two different frequencies alternating in time) are presented. Under the appropriate conditions, the listeners do not report hearing a single sound that alternates in pitch, but rather they report hearing two sounds (two streams) each with its own pulsating pitch.

From Yost (1994)

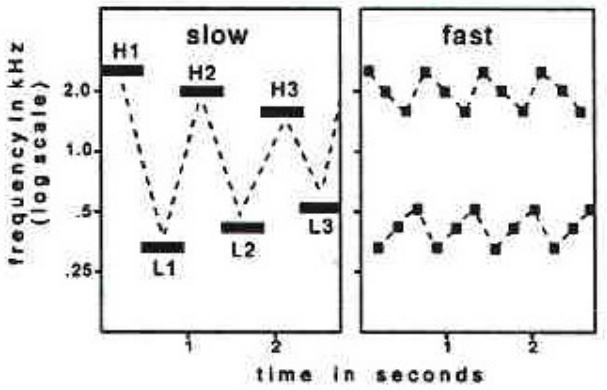
Under some conditions, a single sound source is perceived



But sometimes, two sound sources are perceived

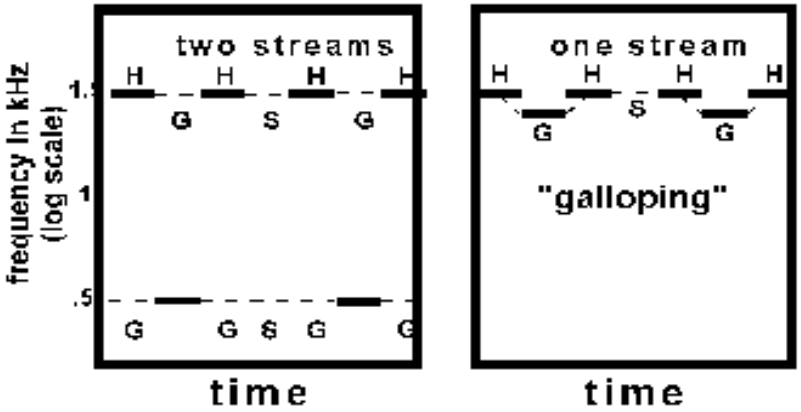


Demo 1:Stream segregation in a cycle of six tones



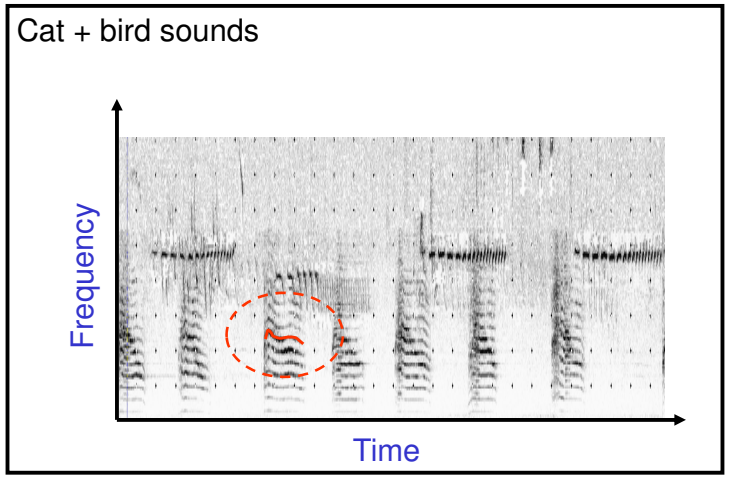
Is it clear what a stream is?

Demo 3: Accelerating galloping patterns, with large and small frequency differences



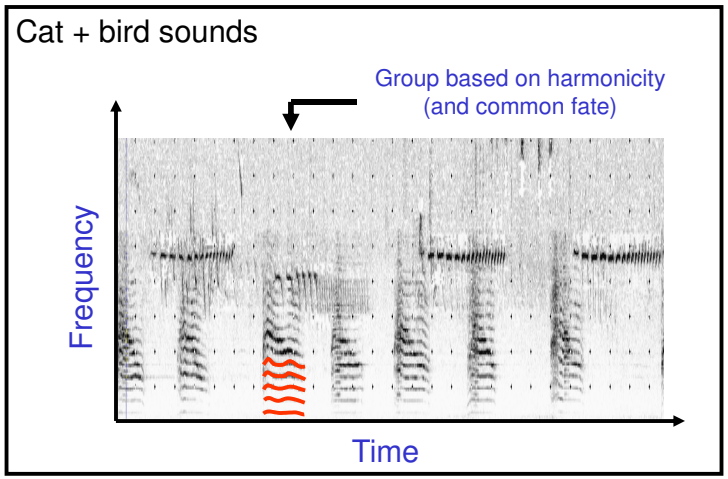
Auditory scenes: continuity

- Group components that are continuous in time or frequency.



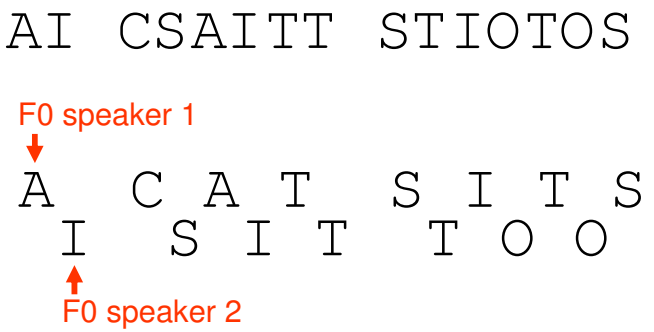
Auditory scenes: similarity

- Group sounds that are similar in pitch, timbre, harmonicity, or location.



Auditory scenes: similarity

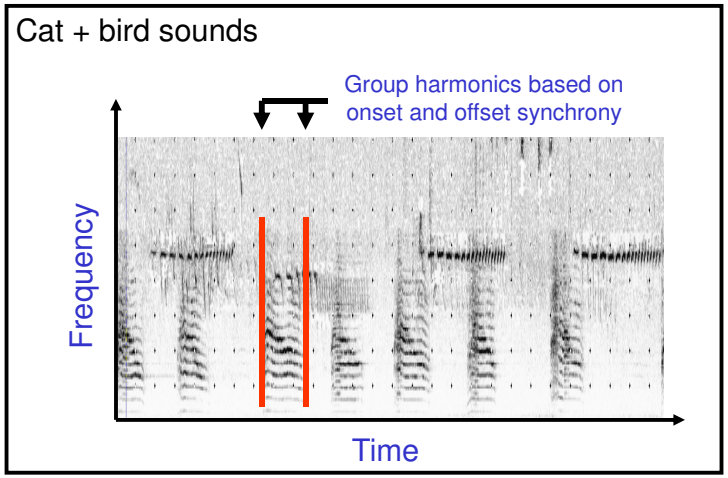
- Can segregate one speaker from another based on differences in voice pitch.
- The harmonics for each speaker can be grouped based on the similarity of their spacing (Assman and Summerfield, 1990).



Letter example from Bregman (1990)

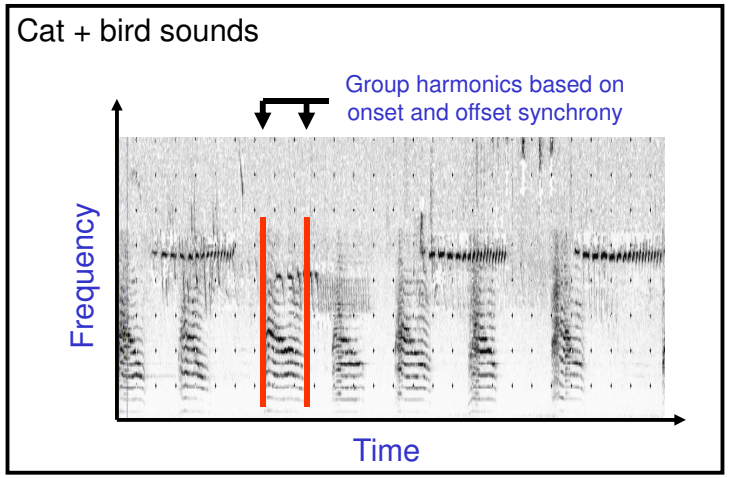
Auditory scenes: common fate

- Group sounds that have a common onset or offset.



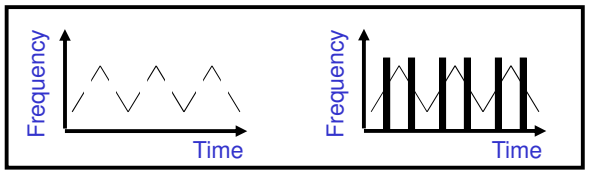
Auditory scenes: common fate

- Group sounds that have a common onset or offset.



Auditory scenes: closure

- We actively use our stored knowledge of sounds to complete segments that have been masked.



Demonstration of gliding tones in background noise

- The auditory system isn't simply filling in the sound with what was there before the noise burst.

http://www.psych.mcgill.ca/labs/auditory/Demo29.html

Gliding tones in background noise

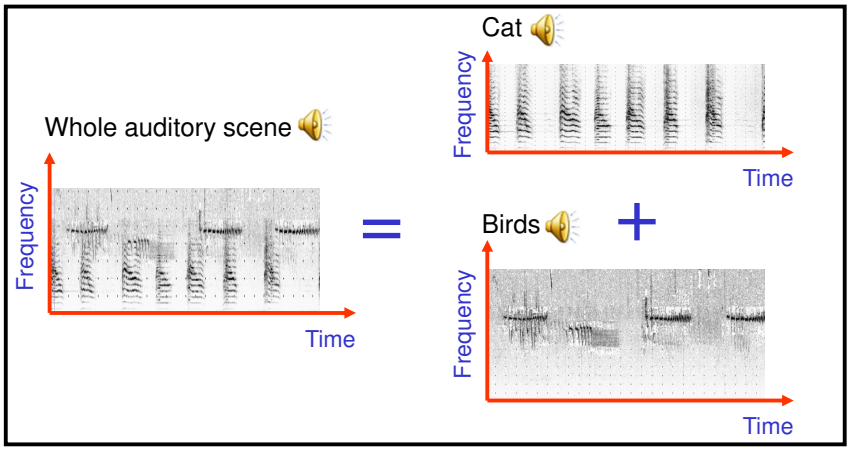
- In this demonstration a rising and falling glide pattern is interrupted by silences or noise bursts.
- During the noise bursts, the glides continue to ascend, descend or change direction.
- The auditory system isn't simply filling in the sound with what was there before the noise burst.

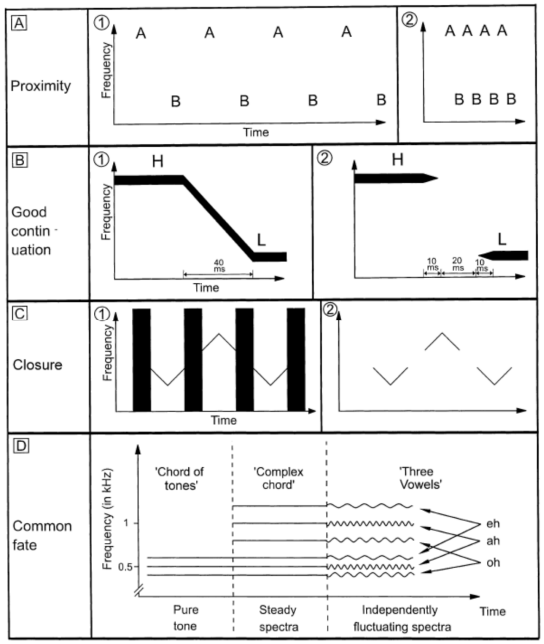
Demonstration of gliding tones in background noise

http://www.psych.mcgill.ca/labs/auditory/Demo29.html

Auditory scene analysis

- Together, these principles enable a separation of the two auditory streams.





Auditory analogues

Purwins *et al.* (2000)