

# Lecture 5: Perception and Action

## Part 4 of 4

# Gestalt Principles and Key Terminology

# Learning Objectives

To provide an introduction to the Gestalt principles of perceptual organisation that apply to both auditory and visual scene analysis.

To provide an explanation of key terminology used in Auditory Scene Analysis and demonstrations of auditory object formation

# Learning Outcomes

To be able to describe the key Gestalt principles

To be able to provide some basic examples of how Gestalt processing applies to both visual and auditory processing of information

To be able to provide a description of the key terminology used in Auditory Scene Analysis and provide examples of auditory object formation

# The Question(s)

A visual scene comprises a number of objects - some in front or, behind one another, some partly hidden, with boundaries between them, with your viewpoint continuously changing. etc.- e.g., busy city.

An auditory scene comprises a number of objects – identifying/following one auditory event alongside other ongoing auditory events: Following a speaker amongst other speakers in a crowded environment with background noise, following one musical line of an orchestra.

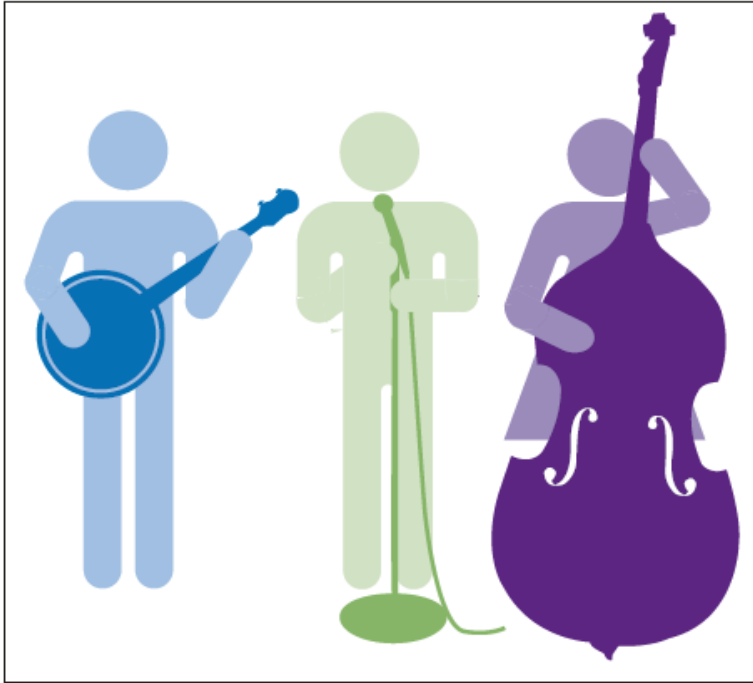
1. How do we manage to do this?
2. How can a machine achieve this?

# Visual Objects and Scenes



Object identification  
despite partial occlusion-  
even when viewed from  
different viewpoints.

# Auditory Objects and Scenes



The **auditory scene** is composed of 3 sounds from a singer, banjo-player and bassist

The auditory information received by the ear is a mixture of these 3 sources



**“Binding Problem”:**  
Segregate sound components that come from different sound sources  
Group sound components that come from the same sound source



Listener hears each source as a distinct **auditory object**.

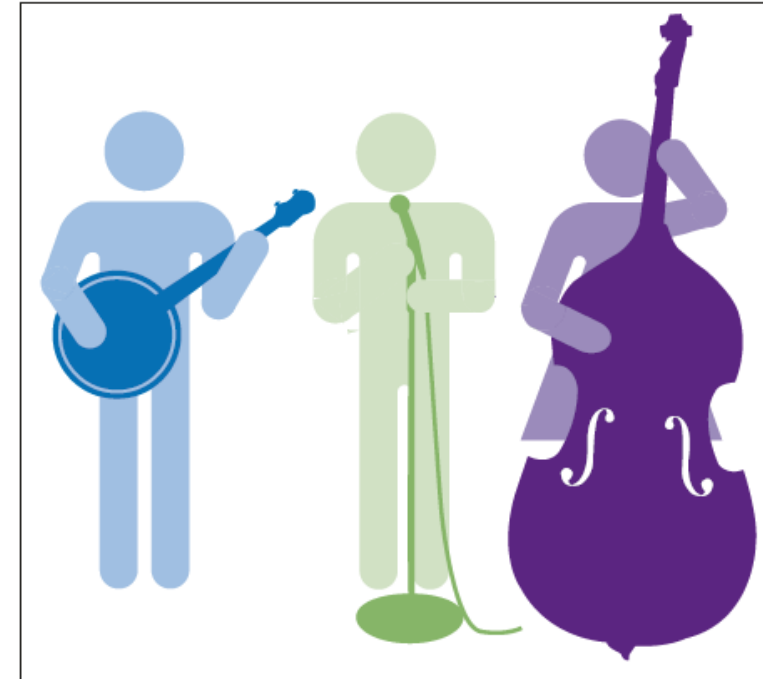
# How can we tell what parts of the Visual/Auditory scene correspond to the different objects?

## Visual Objects and Scenes



Shape, Colour, Location, Texture ...  
Visual system extracts these features & can group appropriate ones together

## Auditory Objects and Scenes





# Gestalt Principles

We perceive the world as stable and composed of discrete objects even though auditory and visual inputs are often ambiguous.

How?

Perceptual organisation:

Gestalt principles govern grouping of elements which can be applied to the auditory or visual scene to create larger objects.

(Gestalt Psychologists working in 1900s)

Aim of the principles: To organise our perceptual world into the simplest pattern consistent with sensory information and experience.

These perceptual principles guide in analysing a complex scene (visual or auditory).

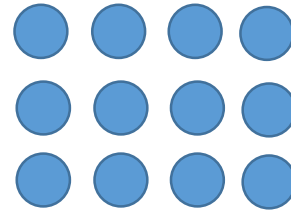
Application of these principles results in a grouping of the elements of a visual/auditory scene that come from the same auditory event/visual object and segregating those elements of a visual/auditory scene that do not come from the same auditory event/visual object.

# Gestalt Principles- Apply to BOTH Visual and Auditory Processing of scenes

Closure



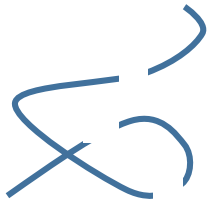
Proximity



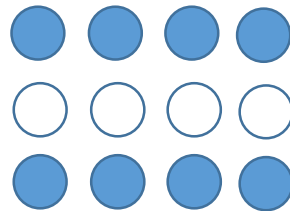
Synchrony



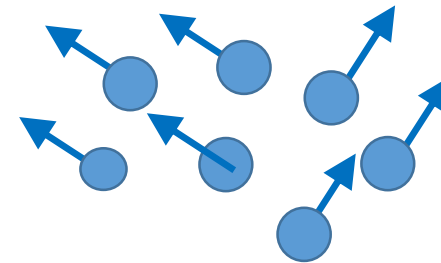
Continuation



Similarity

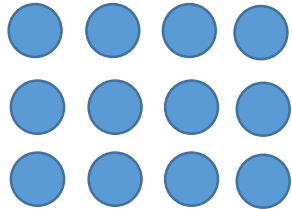


Common Fate

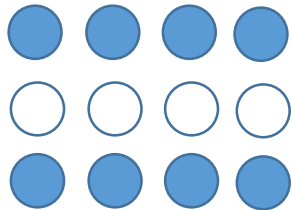


# Similarity

Similar things appear to be grouped together.



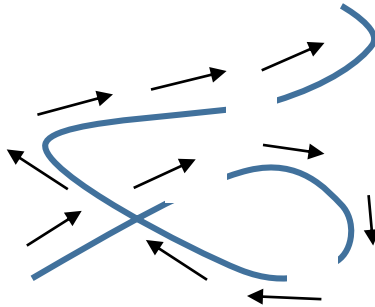
Perceived as horizontal rows,  
vertical columns or both.



Perceived as horizontal rows.

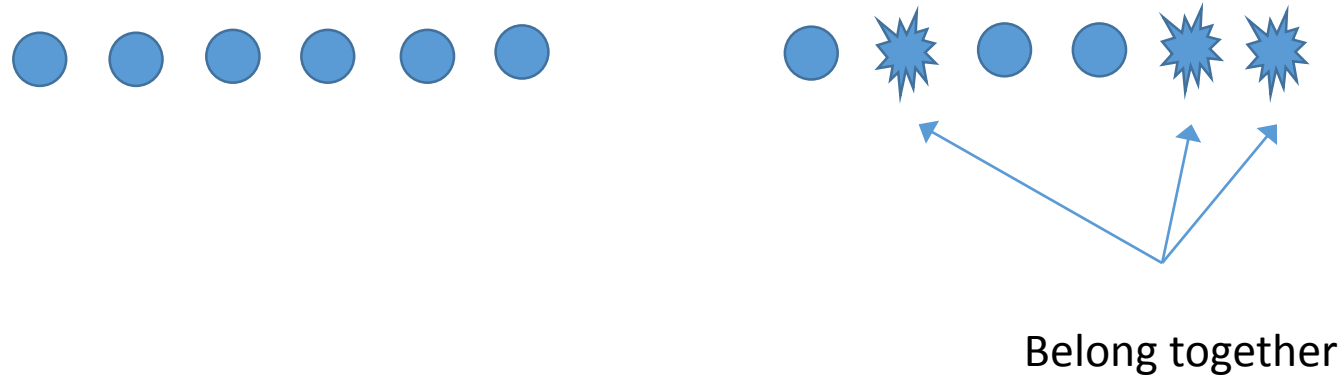
# Continuation

Points that when connected result in straight or smoothly curving lines are seen as belonging together, and the lines tend to be seen in such a way as to follow the smoothest path.



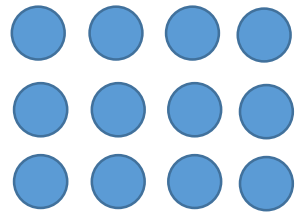
# Synchrony

Events that occur at the same time are perceived as belonging together.

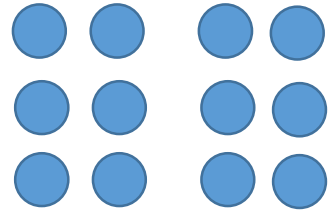


# Proximity

Things that are near each other appear to be grouped together.



1 group

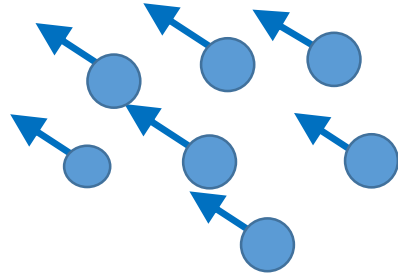


2 groups

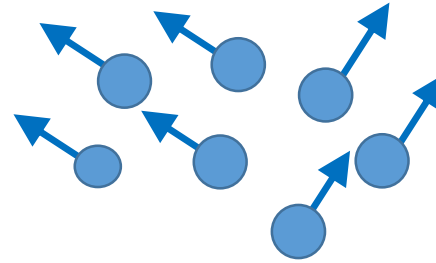
These set of circles appear to be grouped in vertical columns unlike the set of circles on the left

# Common Fate

Things that are moving in the same direction appear to be grouped together.



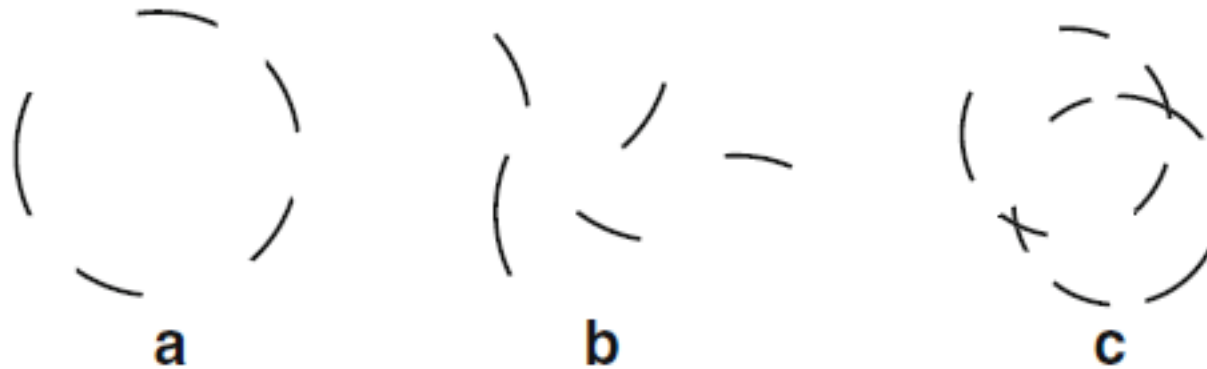
1 group



2 groups

# Closure

Fragmented elements will tend to assemble into a meaningful configuration. We tend to ignore gaps and our mind can fill in the missing information to create familiar shapes and images.



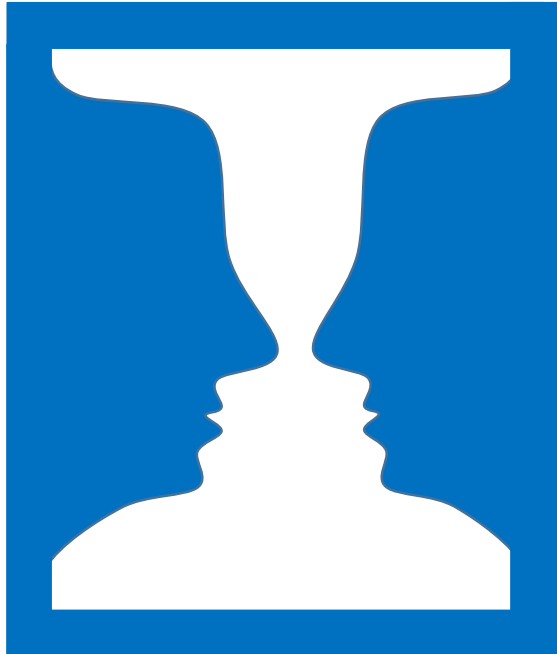
**Fig. 1** **a** A circle formed from fragments via closure; **b** the same fragments but rearranged to prevent closure; **c** fragments from two circles which can be segmented using closure



# Perceptual Segregation

Perceptual segregation of one object from another  
Often also referred to as figure-ground segregation.

A separate object (figure) that stands out from its background (ground).



E.g., reversible figure-ground patterns:

Vase = figure, then blue = background

Faces = figure, then white = background

What factors determine which area is figure? May be symmetry, smaller area, orientation, meaningfulness.

## Auditory Objects and Scenes

(Visual scene analysis will be covered in the next lecture)

# Auditory Objects and Scenes

**Auditory Scene:** The auditory components that form the auditory events around us

**Auditory Scene Analysis:** Auditory system has to make sense of the mixture of auditory components being received

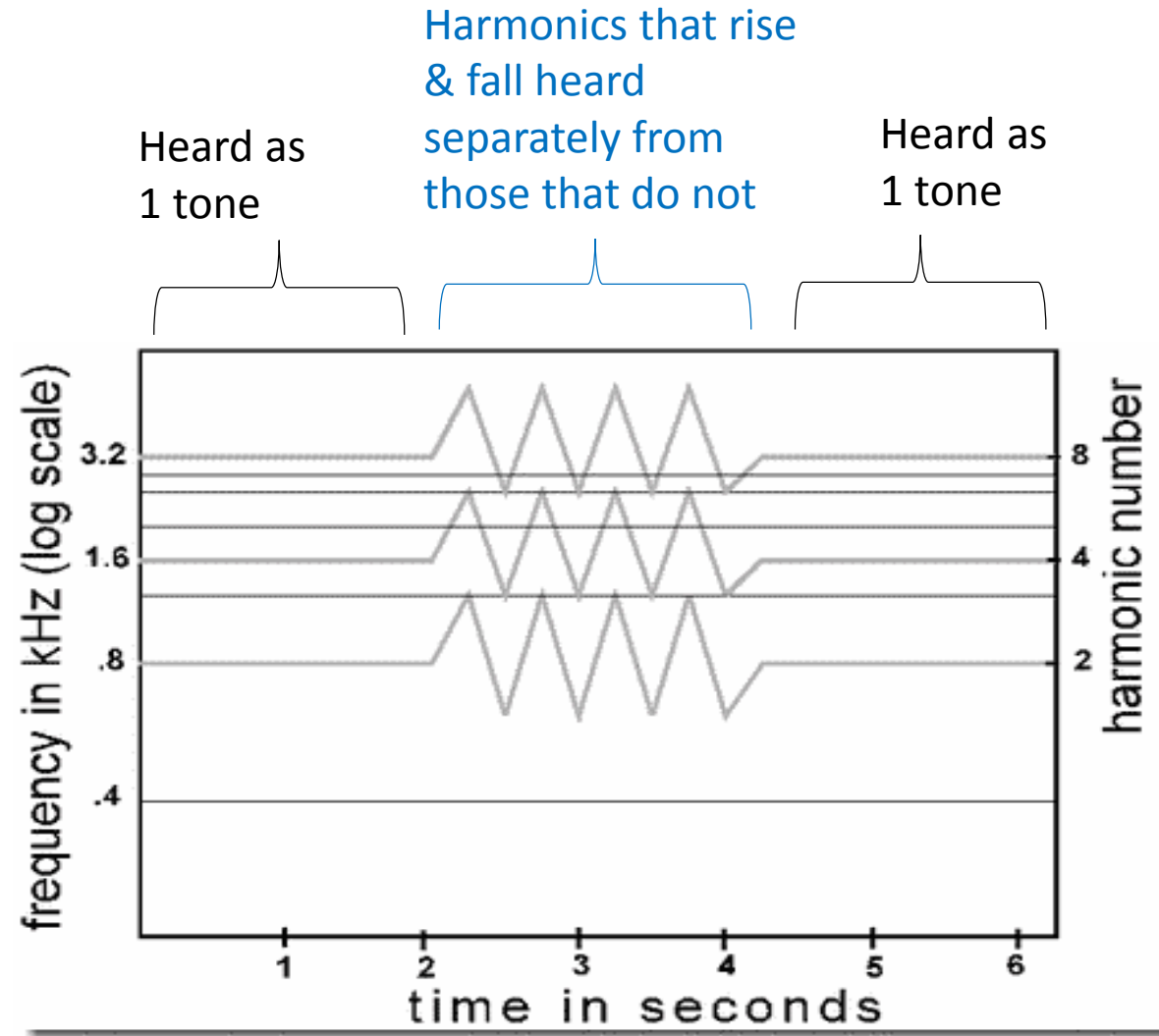
**Computational Auditory Scene Analysis:** Machine systems that can achieve sound-source separation by exploiting perceptual principles

**Group:** Auditory system is able to group together sound components that come from different sound sources

**Auditory stream/stream:** Percept of a group of auditory components forming a coherent whole appearing to come from a single sound source/event

**Auditory Object:** The percept that the auditory components create in order for us to identify the source

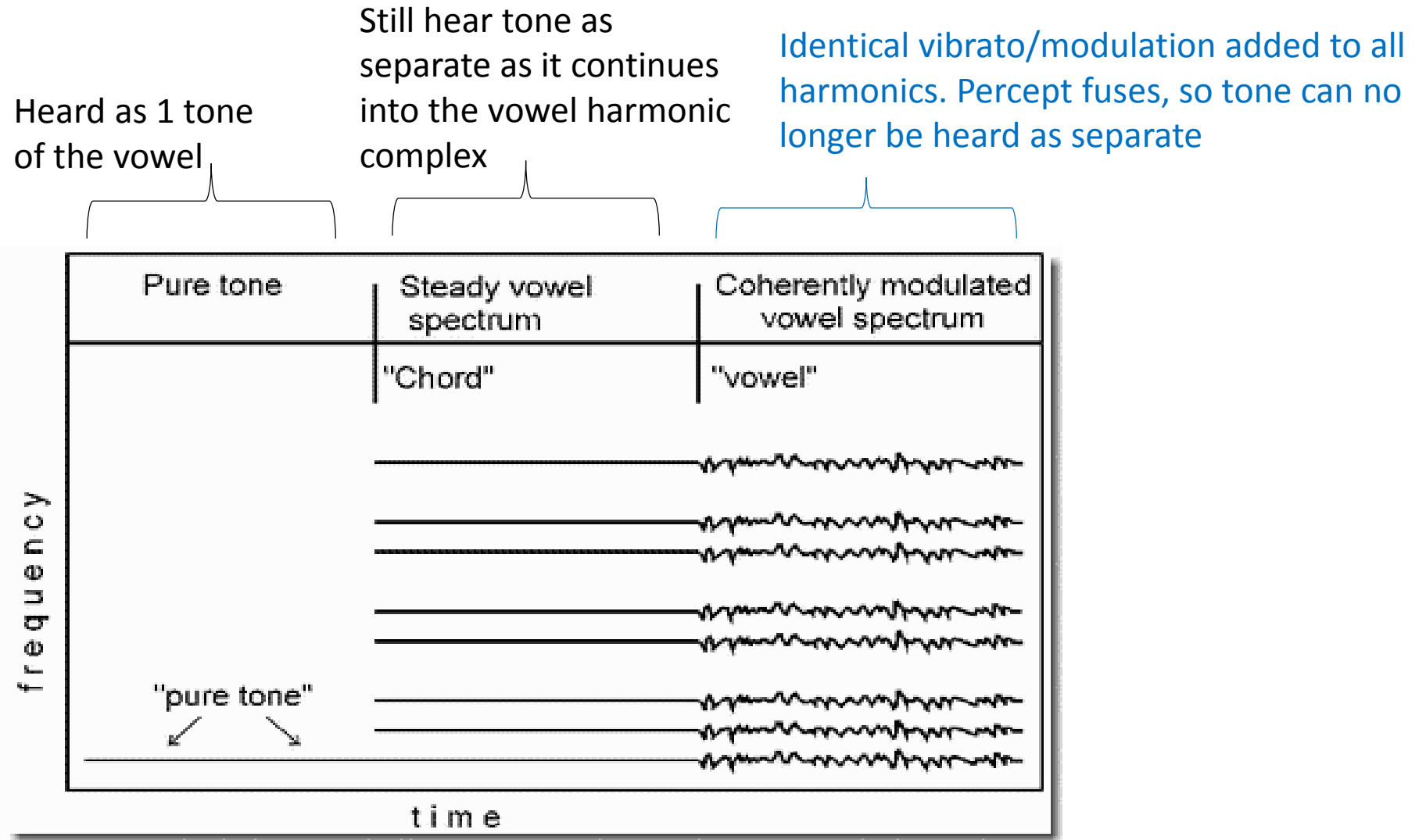
# Fusion by Common Frequency Change



Demo 19 from  
Bregman's  
Auditory Demo  
as cited below

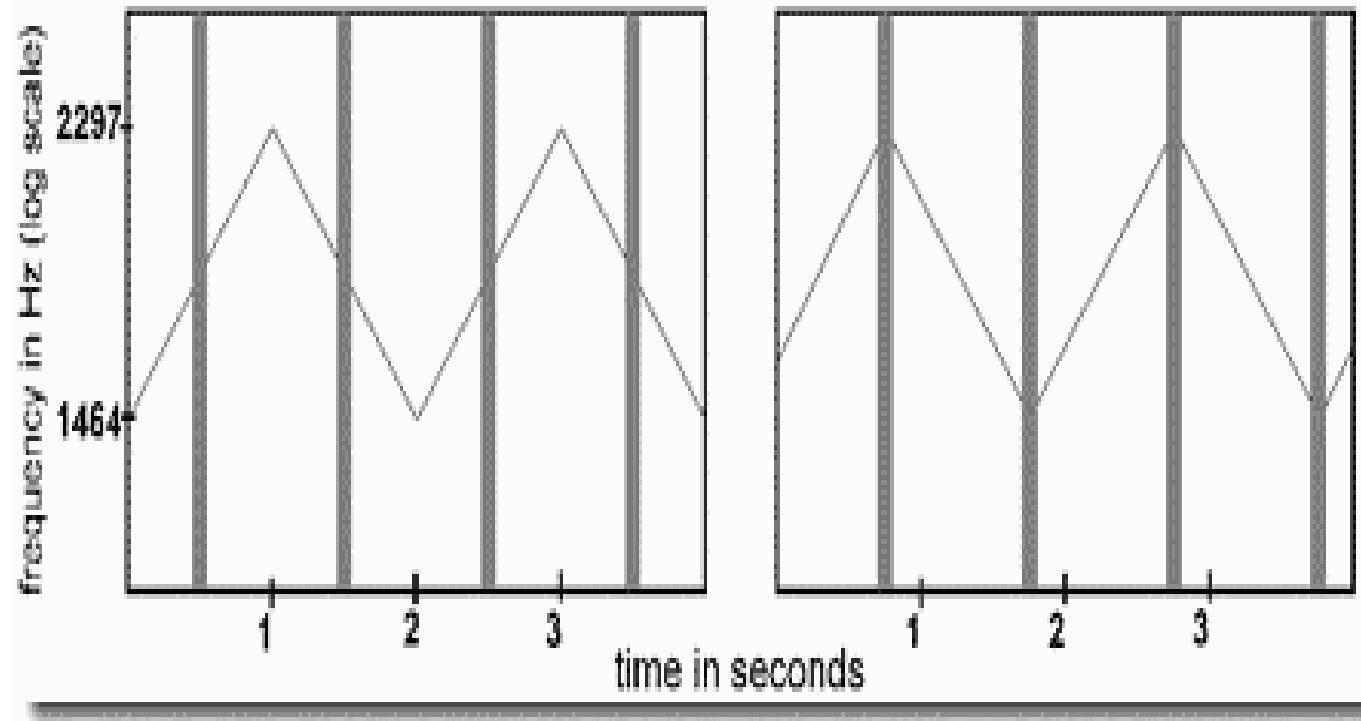
Auditory demo from: <http://webpages.mcgill.ca/staff/Group2/abregm1/web/>  
Bregman, A.S., 1994. *Auditory scene analysis: The perceptual organization of sound*. MIT press.

# Role of Frequency Micromodulation in Voice Perception



Demo 24 from  
Bregman's Auditory  
Demo as cited below

# Perceptual Continuation of a Gliding Tone Through a Noise Burst



Variation of the continuity illusion.

Repeatedly rising and falling glide pattern with interruptions.

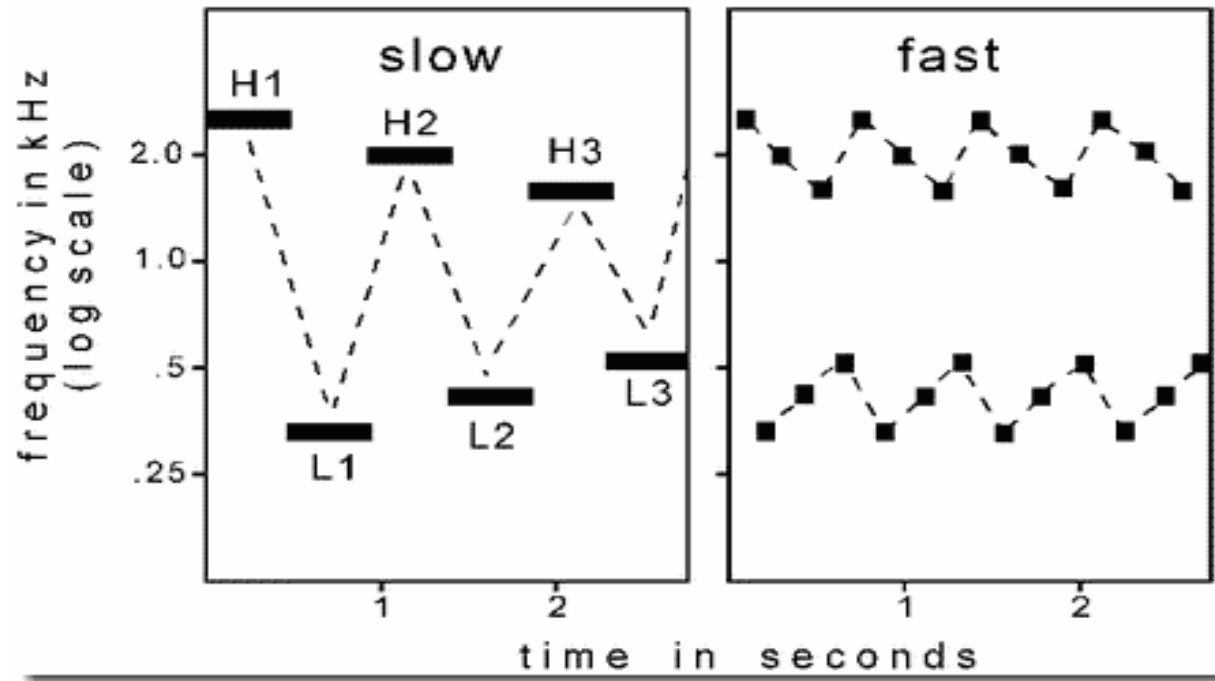
When interruptions are silences- hear short gaps.

When interruptions are noise- hear complete glide pattern.

Demo 29 from  
Bregman's Auditory  
Demo as cited below

# Auditory streaming example: Sequential Integration

Sequential Integration: When parts of a spectrum are connected over time



An example is the connecting of the notes of the same instrument together to create the melody carried by that instrument. Another example is our ability to continue to hear a sound when other sounds join it to form a mixture.

Demo 1 from  
Bregman's  
Auditory Demo  
as cited below

Auditory demo from: <http://webpages.mcgill.ca/staff/Group2/abregm1/web/>  
Bregman, A.S., 1994. *Auditory scene analysis: The perceptual organization of sound*. MIT press.

# Auditory Streaming Example: Segregation of a Melody From Interfering Tones

Telemann's Sonata in C Major, from Der Getreue Musikmeister.



Demo 6 from  
Bregman's Auditory  
Demo as cited below

Formation of perceptual streams in Western music.

Baroque period composers  
(approximately 1600-1750)

Music in which individual instruments rapidly alternated between a high and a low register.

-> two intertwined melodic lines ("virtual polyphony" or "compound melodic line").

-> Some degree of segregation.



# Auditory Streaming Example: Repeating Cycle of Notes with Pitch Difference

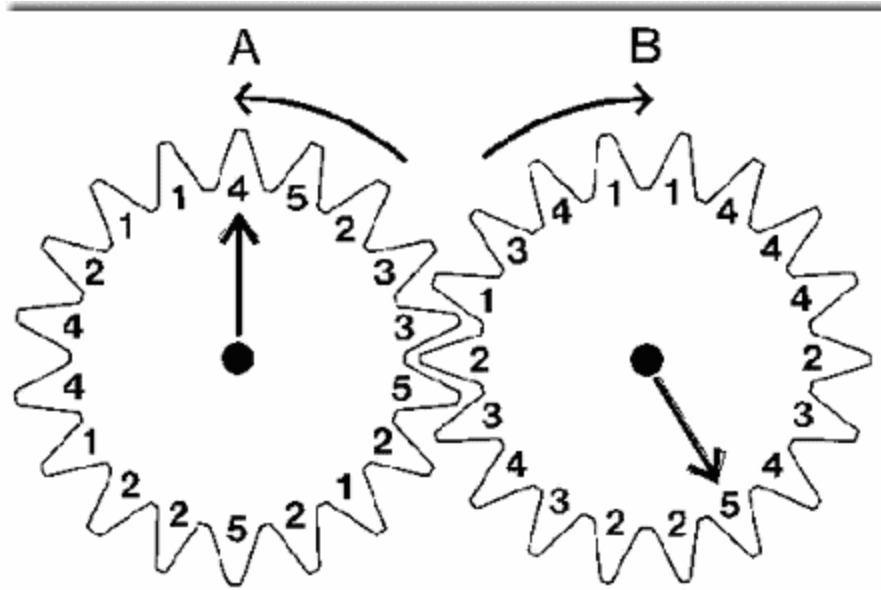
Similar to previous example.

"SSematimba ne Kikwabanga", from Buganda.

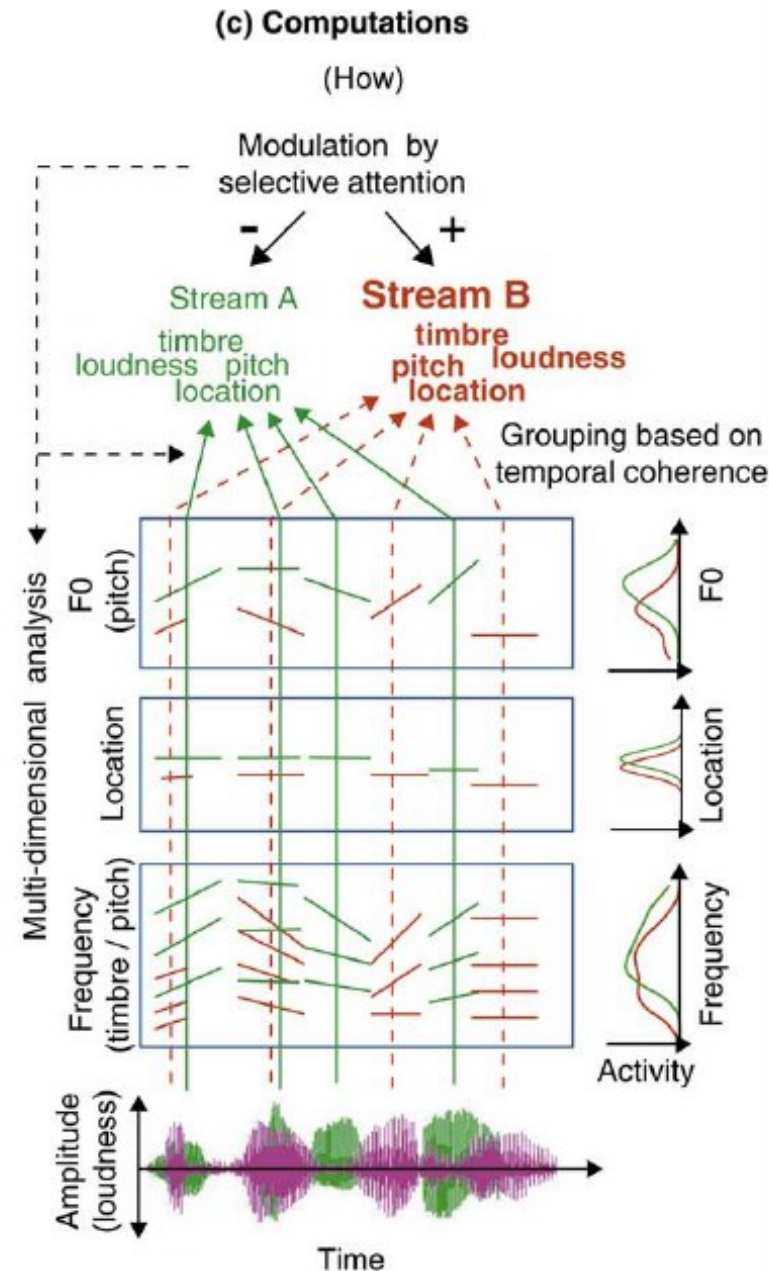
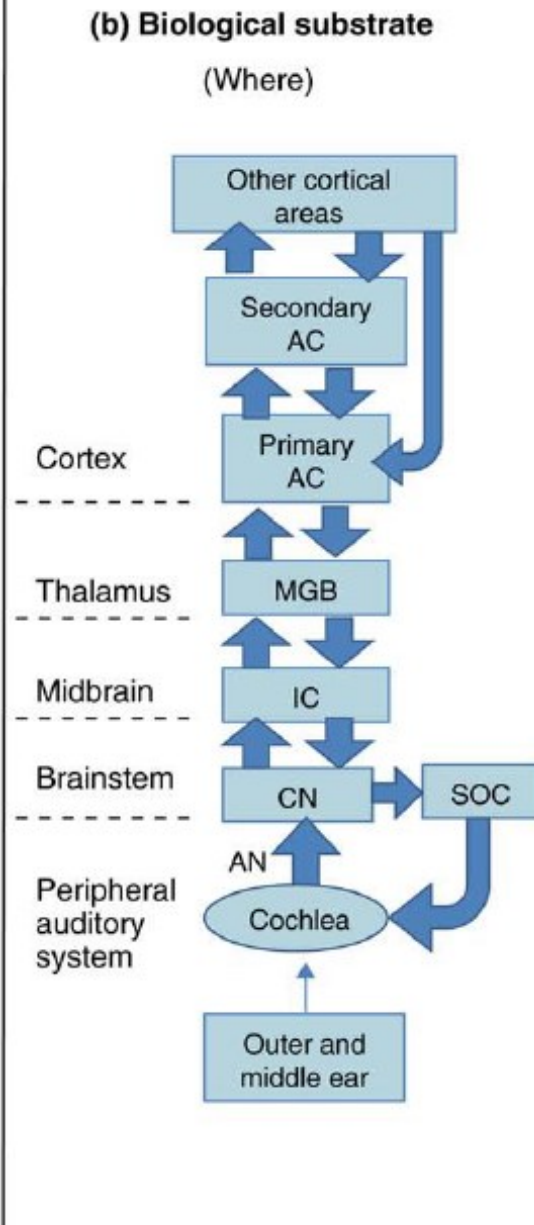
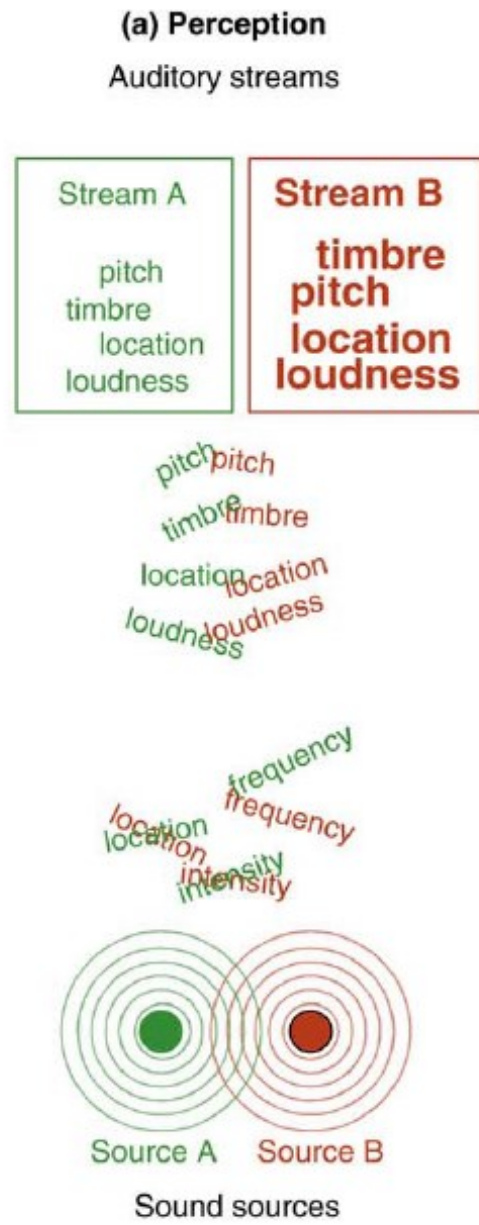
Each of two players plays a repeating cycle of isochronous notes (notes equally spaced in time), the notes of each player interleaved with those of the other.

Interleaving of the two cycles creates high and low separate perceptual streams with irregular rhythms.

Style is typical of music for the amadinda (a twelve-tone xylophone).



Demo 7 from Bregman's  
Auditory Demo as cited  
below



Psychophysical, biological, and computational aspects of the auditory scene analysis problem.

Contribution of top-down attentional components on Auditory scene analysis.

# Overall Summary

Described the concept of auditory scene analysis.

Gestalt principles that apply to both visual and auditory processing

Provided descriptions and examples of some Auditory demonstrations of grouping to form auditory objects

# Resources

## Essential:

Sensation and Perception- E. Bruce Goldstein: Chapter “Perceiving Objects and Scenes”. [Chapter 5 in 8<sup>th</sup> Edition] and Chapter “Sound Localization and the Auditory Scene” (Chapter 12 in 8<sup>th</sup> Edition].

Auditory demo and explanation of stimuli from: <http://webpages.mcgill.ca/staff/Group2/abregm1/web/>