Auditory Scene Analysis

Dr. Maria Chait, UCL Ear Institute
Expected learning outcomes:

• Understand the tasks faced by the auditory system during everyday listening.

• Know the major Gestalt principles.

• Understand the major principles of ‘auditory scene analysis’.
Suggested reading
Auditory Scene Analysis

• 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.
Auditory Scene

Cat

Birds
Mixture of two speakers:
(“binding problem”)

- **segregate** the components of the sound that come from different sound sources.
- **group** the components of the sound that come from the same sound source.
Auditory Scene Analysis

- 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**.

(slide from: Stuart Rosen)
Similar problem in Vision?
The principles of auditory scene analysis are similar to those for visual scenes.

How do we know what parts of the visual scene correspond to different objects?

How do we know what parts of a visual scene correspond to different objects?
• The principles of auditory scene analysis are similar to those for visual scenes.
• How do we know what parts of a visual scene correspond to different objects?
Features that make up visual objects

- Color
- Shape
- Location
- Texture
- ....

The system extracts these features then decides which ones group together.
Visual Scene Analysis
Main principles proposed by *Gestalt* psychologists (*gestalt = form or pattern*) in the early 20th century.

A set of *Gestalt grouping rules* that describe which elements in an image belong together to form an object.

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

(slide from: Stuart Rosen)
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

(slide from: Stuart Rosen)
Visual grouping by similarity

- Similar things are perceived as one group.

Perceived as 6 vertical objects

Perceived as 6 horizontal objects

(slide from: Stuart Rosen)
Visual completion by closure

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

Perceived as an obscured disc

Not…

(slide from: Stuart Rosen)
Visual completion by closure

• Perception is a constructive process – an interaction of stored knowledge and incoming sensory information.

We perceive a pentagon. Our mind fills in familiar shapes.

(slide from: Stuart Rosen)
Visual completion by closure

- A whole cat and not disconnected shapes

(slide from: Stuart Rosen)
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
(slide from: Stuart Rosen)
Visual grouping by proximity

- Things close together are perceived as one group.

(slide from: Stuart Rosen)
Visual grouping by continuity

- Lines are seen as following the smoothest path.

 Seen as...

 Not...

(slide from: Stuart Rosen)
Visual grouping by common fate

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

(slide from: Stuart Rosen)
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
Features that make up Auditory objects

- Pitch
- Timbre
- Loudness
- Location

The system extracts these features then decides which ones group together.
Mixture of two speakers:

In order to be able to follow the speech signal we need to integrate within time (across frequency) and across time.
Across-frequency grouping
Fusion based on common Frequency modulation:
Gestalt principles

- Proximity
- Similarity
- Continuity
- Closure
- Common fate
- Disjoint allocation
  - An element of a visual scene must belong to a single object.
- Figure/ground
<table>
<thead>
<tr>
<th>Pure tone</th>
<th>Steady vowel spectrum</th>
<th>Coherently modulated vowel spectrum</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>&quot;Chord&quot;</td>
<td>&quot;vowel&quot;</td>
</tr>
</tbody>
</table>

**Diagram:**
- **Frequency**
- **Time**

**Legend:**
- "pure tone"
Mixture of two speakers:

In order to be able to follow the speech signal we need to integrate within time and across time.
Across-time grouping
The continuity Illusion

Vision
Gestalt principles

- Proximity
- Similarity
- Continuity
- Closure
- Common fate
- Disjoint allocation
  - An element of a visual scene must belong to a single object.
- Figure/ground
Another example of continuity Illusion (picket fence effect in speech)
Auditory scenes: closure

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

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

http://webpages.mcgill.ca/staff/Group2/abregm1/web/downloadstoc.htm#29
Gestalt principles

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

Frequency

Time

A
https://mustelid.physiol.ox.ac.uk/drupal/?q=topics/streaming-galloping-rhythm-paradigm
Gestalt principles

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

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

(slideshow from: Stuart Rosen)
Visual information aids auditory segregation
Audio-Visual binding

Making sense of the Acoustic environment
Bi-Stable auditory perception
Bi-stable perception
Bi-Stable auditory perception

A

B

Time
We learn the world by forming internal models of the expected behaviour of sources in the environment and testing those against the input.

In certain cases the input is consistent with 2 interpretations (that are mutually exclusive) and we then randomly shift between them.
Expected learning outcomes:

• Understand the tasks faced by the auditory system during everyday listening.

• Know the major Gestalt principles.

• Understand the major principles of ‘auditory scene analysis’.