

# Auditory Scene Analysis

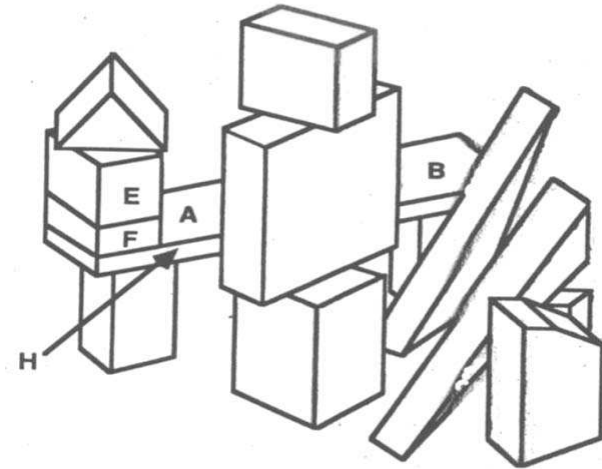
*Otherwise known as*

Auditory Grouping  
Auditory Streaming  
Sound source segregation

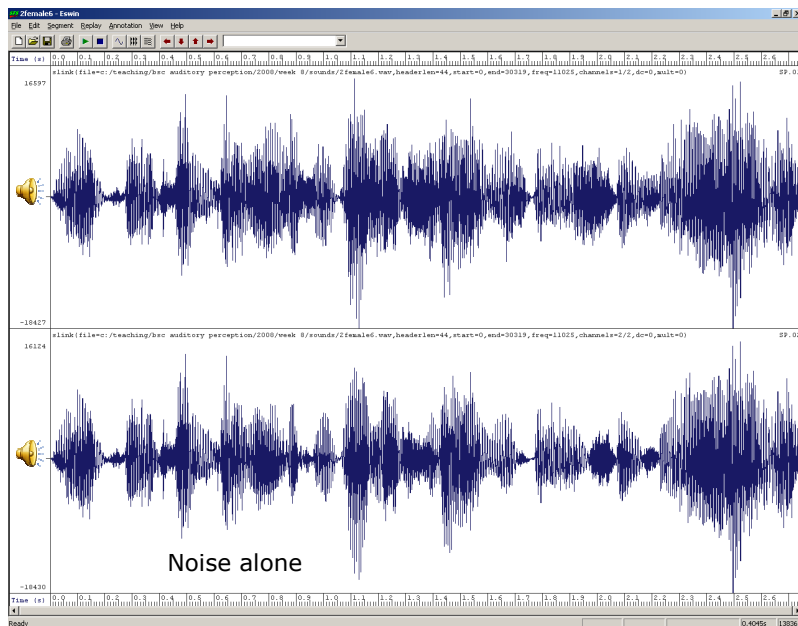
Assigning acoustic/auditory features to  
distinct objects or sources of sound

1

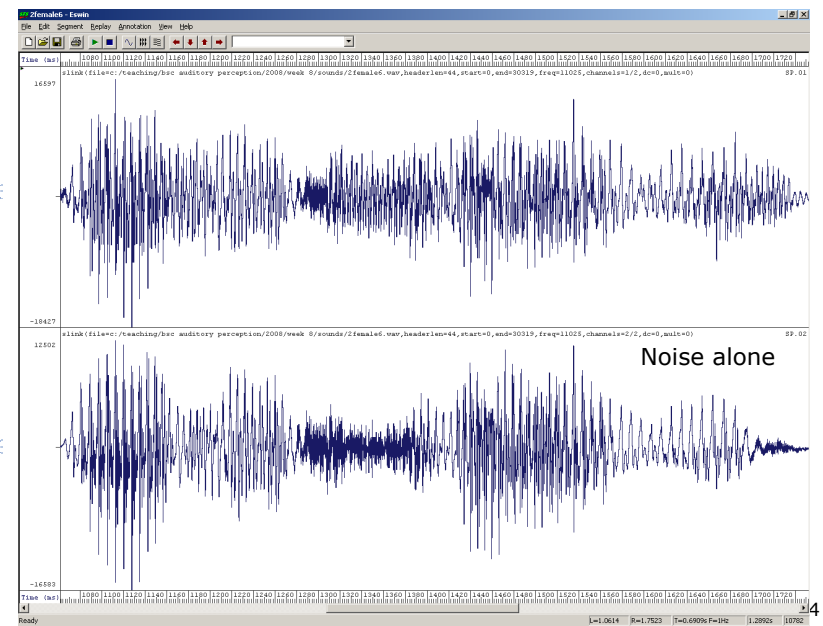
The visual analogue: Assigning  
visual features to distinct objects



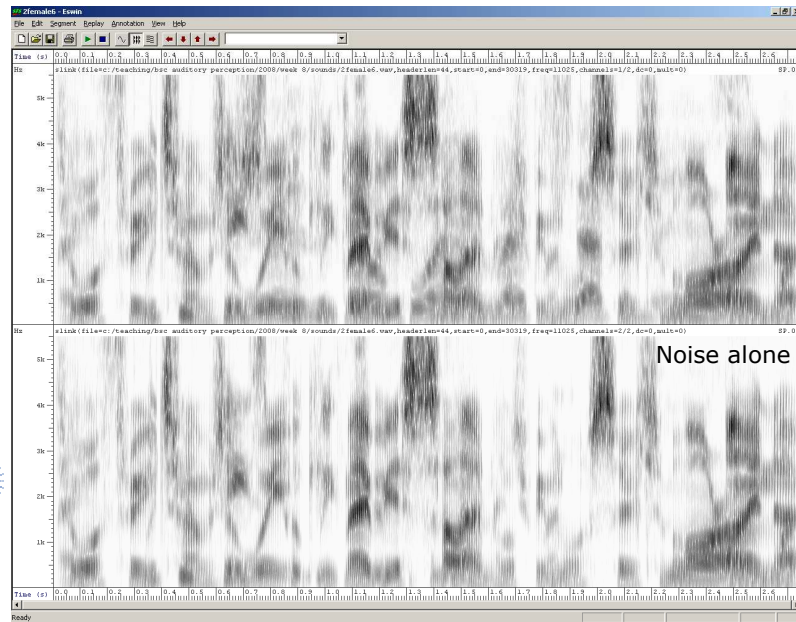
2



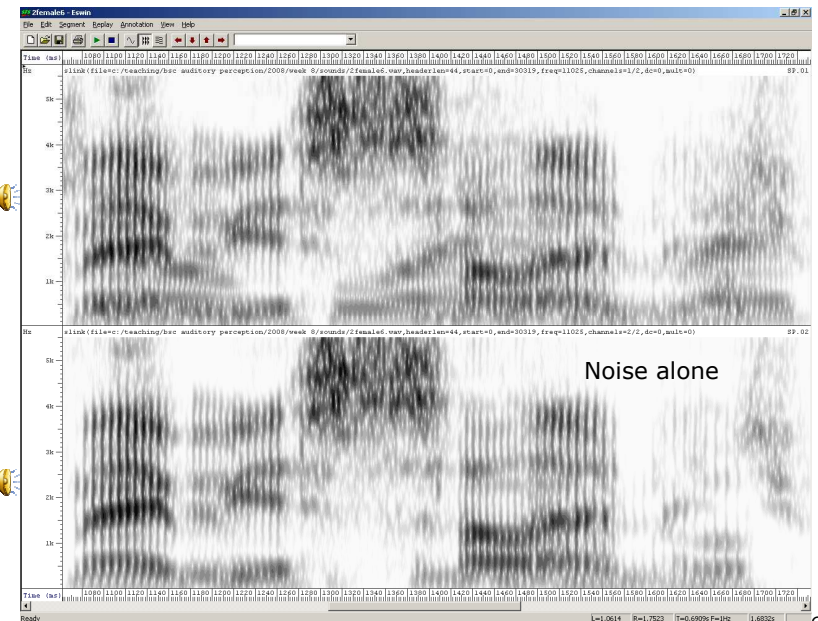
3



4



5



6

## Visual examples of Gestalt principles



### Law of Pragnanz:

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.

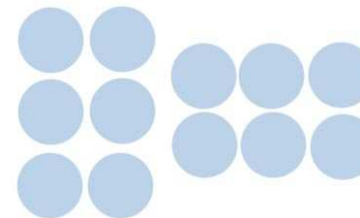


### Law of Similarity:

Items that are similar tend to be grouped together.

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

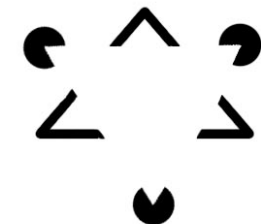
## Visual examples of Gestalt principles



### Law of Proximity:

Objects near each other tend to be grouped together.

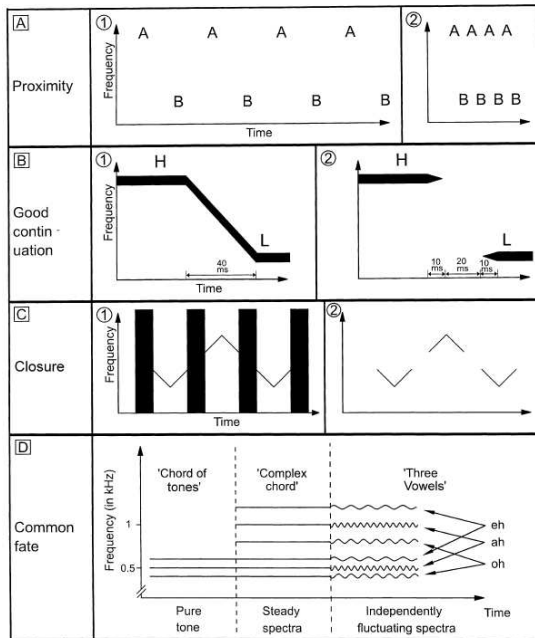
The circles on the left appear to be grouped in vertical columns, while those on the right appear to be grouped in horizontal rows.



### 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.



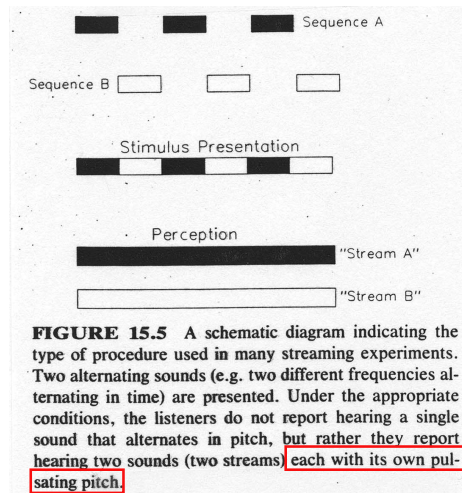
## Auditory analogues

Purwins *et al.* (2000)

## Sequential segregation

10

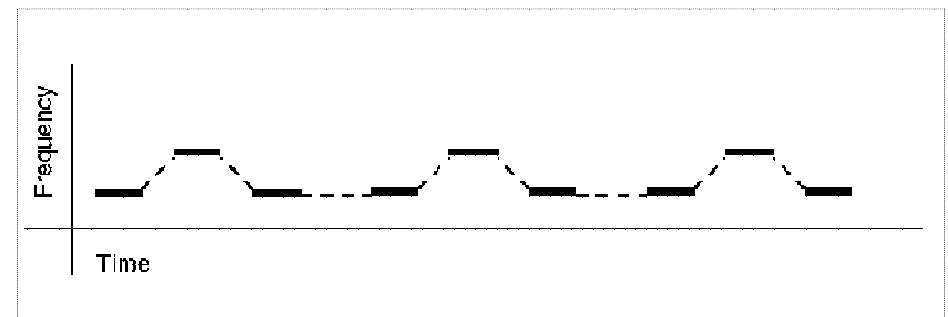
## Auditory streaming



From Yost (1994)

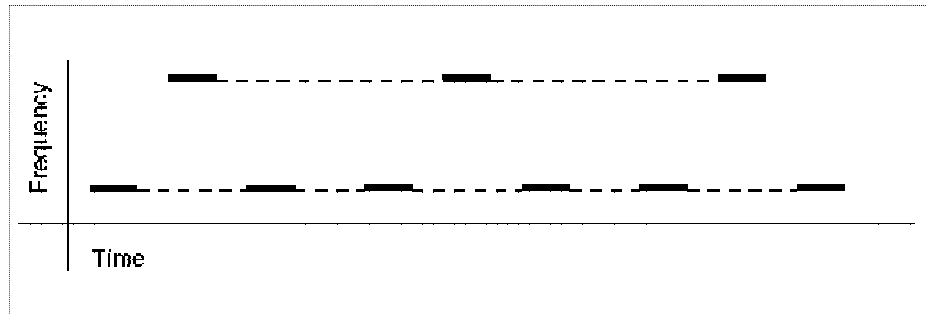
11

## Under some conditions, a single sound source is perceived



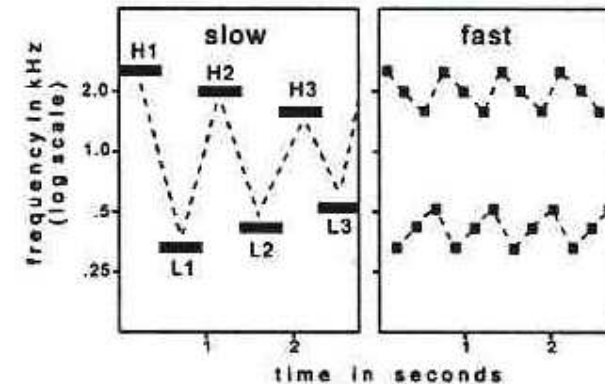
12

But sometimes, two sound sources are perceived



13

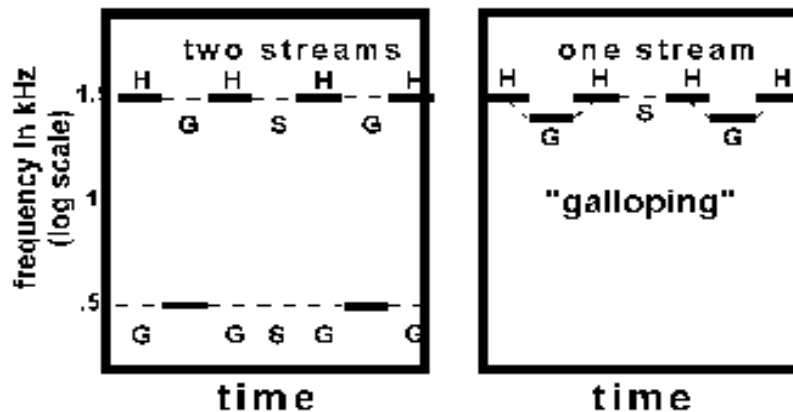
Demo 1: Stream segregation in a cycle of six tones



Is it clear what a stream is?

14

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



15

Old-plus-new heuristic

- A method for decomposing mixtures of sounds
- When a spectrum becomes more complex, interpret this as ...
  - a continuing old sound
  - joined by a new one supplying the additional components
- So the interpretation of a particular sound (in terms of assigning sources) will be affected by its temporal context (things that just happened)

16