How Does the Hearing System Perform Auditory Scene Analysis?

Georg M. Klump
Contents

Key question

Basic auditory system

2 ways of processing

Segregation of sequential tone pulses

Object formation by common modulation

Object formation by spatial processing

Good continuation: top-down processing

Conclusion

Future study
Key questions

How does the brain perform the auditory scene analysis?

Auditory Scene Analysis
- Termed by Albert Bregman
- The process by which the human auditory system organizes sound into perceptually meaningful elements.

Auditory Stream
- A perceived melodic line
Basic auditory system

Cochlea as a **Fourier transformer**
2 ways of processing

**Bottom-up processing**

- Segregation of sequential tone pulses
- Object formation by common modulation
  - CMR, MDI
- Object formation by spatial processing

**Top-down processing**

- Perceptual restoration of interrupted speech
Segregation of sequential tone pulses

The segregation of tone sequences depending on the **proximity of the tones in time and frequency**

- Percepts as one stream like galloping rhythm or separated stream

Open circle: Human perceptual data
Filled circle: European starling perceptual data
Filled square: L2 in the starling forebrain
Segregation of sequential tone pulses

Responses in the primary auditory cortex of the macaque monkey

Segregation of sequential tone pulses

Fishman thought about local inhibitory circuit in the primary auditory cortex about the suppression. But this can be the result of across-frequency processing at lower levels (cochlear nuleus)

Build-up effect (Micheyl et al.) : might be affected by attention

Other cues : envelope, interaural time, intensity..
Object formation by common modulation

**Temporally coherent changes** (correlated variation) of each frequency: potent cues of grouping of the components

**CMR (Correlated Modulation Release)**

CMR shows us that if the difference frequency band signals have (temporally) correlated amplitude fluctuations, the auditory system can know that (its’ threshold value is decreased)

1. Analysis of the excitation within a single frequency channel centered on the signal (Human)
2. A comparison of excitation across different frequency channels (other animals)

**MDI (Modulation detection interference)**

In MDI also means the interaction of multiple frequency channel
Object formation by common modulation

For detect to target signal, Best efficient S/N filter(s) is used

If target signal and coherent modulated noises are presented simultaneously, the threshold of target signal is decreased.

Release value is bigger when fluctuation is slow, and noise bandwidth is broad
Object formation by spatial processing

To identify different locations of sound source, about duration of 100ms is needed

a model allowing representation of two sound sources in the ICx of barn owl when presented with temporally uncorrelated noise bursts (Takahashi, Keller (1994))
Good Continuation : top-down processing

The perceptual restoration of interrupted speech could be an evidence for top-down processing.

Humans, birds could perceive one continuous stream.

Sound perception (speech restoration) could be supported by memory process and perceptual learning, if the appropriate excitation is present.

Conclusion

**Auditory scene analysis**
through combination of multiple stimulus dimensions like location, timbre, amplitude...

And **distributed population coding** is able to combine these information
(Each neural excitation encodes relevant stimulus characteristics)

Analysis is performed through top-down process and bottom-up process both
Future study

Distributed populations of cortical neuron in Visual scene analysis

Temporal correlations of distributed neural activity
Temporal patterns of the neural response using stimulus paradigms