Lecture 3

Rhythm and Scene Analysis
(adapted from Mari Jones)
I. Scene Analysis: A Neo Gestalt view
What’s a Scene?

The term *scene* is taken from vision. Two cases are most common:

1. A stationary viewer
2. A moving viewer

The scene is the static visual layout.
Sound Scenes

A common assumption:

We use the proximal sound stimulus to recover the location and/or identity of a distal, vibrating, object within a scene.

This is the source identity principle.
Bregman’s Theory of Scene Analysis

Bregman builds on this idea:

Goal of auditory perception…
is identification of one or more sources of a proximal mixture of sounds.

*If an unfolding sound pattern is perceived as a structurally intact*,
i.e. as *one stream*,
*then it must come from ONE source.*
Other patterns?

**Patterns NOT perceived as intact…**

*If a pattern is perceived to break up, i.e. form N different subjective units, streams,*

*then, it must ...

*come from N separate sources.*
Example: A sound pattern

- Time
- high
- low
- frequency

A stream? One source?
Another source?
The Big Question

What makes a pattern seem structurally intact?

Or... What determines a stream?
Bregman’s Answer

Involves a two stage theory

in which

Neo Gestalt principles are critical.
The Neo Gestalt View

Two Stages

I. Primitive processes
   Innate, automatic, effortless,
   Determine coherent perceptual groups at fast rates.

II. Schema driven processes
   Learned, voluntary, effortful,
   Determine attention to groupings at slow rates.
Auditory Scene Perception

Stage I: Primitive Processes
Operate on Gestalt grouping principles:

- proximity in frequency and time;
- similarity (e.g., timbre);
- continuity;
- exclusive allocation… and so forth

These are domain general

(\textit{i.e.} speech, music, natural environs, etc.)
Proximity vs Continuity

Tougas and Bregman (1985)

In a series of famous studies pitted proximity against continuity.

Goal: To assess which principle dominated, and to dispute a trajectory explanation.

Bouncing versus Crossing trajectories
Stream segregation of high and low bands of noise.
How Frequency affects Streaming
Rhythm Perception

A stream usually conveys a rhythm

Speeding up a sequence can break it, perceptually, into interleaved streams

This changes its perceived rhythm!
Rhythm

Rhythm perception… depends on the *relationship* of frequency change ($\Delta f$) to time change ($\Delta t$).

*This slow melody…*

*is perceived to have a ‘galloping‘ rhythm*
But... That Rhythm Disappears

In a **faster** melody with the same tones.

Now we hear two melodies (streams) each with an **isochronous rhythm.**
A Neo Gestalt Explanation

Builds on Gestalt proximity:

1. **Temporal proximity** is heightened at fast rates.
2. But **frequency proximity** dominates.
3. **Two streams** form based on pitch proximity:
   - A high tone stream and a low tone stream.
II. Scene Analysis: A Dynamic Perspective
Auditory Patterns

Most of our environment comprises sources that turn on and off and modulate over time to project pattern structure that is...

motion-like and rhythmic
Important Pattern Properties

1. Motionlike properties: Rates of change.
   
   \[ e.g., \frac{\Delta f}{\Delta t} \]

2. Rhythmic properties: Relative timing.
   
   \[ e.g., \frac{\Delta t_i}{\Delta t_k} \]
Alternative View:
A focus on time

How do rhythmic properties of a pattern affect real time responding to it?
General Assumptions

**Environment**

- Scenes often project dynamic patterns with relative timing that is quasi-periodic.

**Living things**

- Animate creatures possess internal periodicities that can synchronize to patterns.

**Connecting through time**

- Pattern timing determines how living things connect with a changing scene.
Some periodic patterns

Great Gray Owl $T = 700$ ms

Horned Lark

Cardinal
Internal Periodicities:

Biological Oscillations

Many internal periodicities exist; periods of 1 ms, several minutes, years !.

Most studied are circadian rhythms; oscillatory periods of ca. 24 hours.
Connecting

“How do internal oscillations of an organism connect with pattern timing? “

One answer: By synchronizing.

More precisely: By entrainment
Biological Entrainment


- Daan et al. (2001) recently updated this two oscillator model.
Shorter Periods?

In species with refined use of fast sound patterns, Can we postulate, internal oscillations with periods less than the circadian period?
Generalizing…..

Most Biological Oscillators

\(\text{Long natural periods}\)

Respond to \(\text{light changes}\)

Alternatively Other Oscillators

\(\text{Short natural periods}\)

Respond to \(\text{sound changes}\)
Background:

*Rhythmic (periodic) Expectancies*

Quasi-periodicities in Rhythmical sound patterns Induce *Periodic expectancies*

*Perhaps such expectancies reflect entrainments to patterns of stimulus IOIs?*
Back to Synchrony:
A simple premise

When we attend, an internal activity is elicited ... that co-occurs with the attended object.

This activity involves Internal oscillations

(Jones, 1976).
Attending to Dynamic Patterns

Where elements appear… then disappear

\[\text{Time} \rightarrow\]

Attention must occur in synchrony with elements.
Dynamic Attending

*Postulates that:*

- Selective attending in time is paced by stimulus timing.

- Attentional focus in time becomes phase-locked to elements.
III. The Changing Scene: 
*Tracking sounds in real time*
Attending oscillates in time

Resources are periodically distributed as **internal oscillations**.

Inter-onset-interval, IOI

*Pulses of attention energy*
Formalized as an Oscillator Model:

by Ed Large (Large & Jones, 1999)

Stimulus timing paces the oscillator. Oscillator period and phase determine this entrainment.
The adaptive oscillator

Period and Phase change: They draw attending rhythms to “fit” a time pattern.
The entrained oscillator gravitates to an attractor state of:

1. **Phase Synchrony:**
   A zero time difference between an sound onset and a pulse peak.

2. **Period matching:**
   A zero time difference between an IOI and oscillator period (limit cycle).
Theoretical Phase Attractor: Synchrony 

Phase Response Curve, PRC

Relative Phase, $\phi_i$

Phase Correction, $-\phi_x$

{ $\phi_i = 1.0$ }

Entrainment Region
What does this approach purchase?

? A few experiments....
Preview: Our Experimental Designs

Will involve a context rhythm, given by tones:

Followed by a comparison, C, (of some sort)

Time Judgment
Comparison IOI

Pitch Judgment
Comparison tone
Time Judgment Tasks

*Rationale*

If relative timing in sound arrays entrains attending, then the context rhythm should:

- Shape temporal expectancies
- Affect time judgments.
Example Study:

Judgment Task: Is a comparison time interval “shorter”, “same” or “longer”?

“Ignore context IOIs“

Context IOIs are 600 ms.
This rhythm should synchronize attending.
Standard Ending Varies:

Expected Standard Ending: \( = 600 \text{ ms} \)

Unexpected Standard Endings:

- Early = 579 ms
- Late = 621 ms
- Very Early = 524 ms
- Very Late = 676 ms

A comparison IOI is always yoked to standard IOI
When a standard ends as expected, the oscillator’s period is accurately preserved…as a memory of the standard IOI.
The Reason: The Attentional Pulse

Expected time

Pulse location

Pulse location depends on rhythmic context
One Prediction:

A quadratic accuracy *Expectancy Profile*:

![Graph showing a quadratic accuracy expectancy profile with three labeled points: Very Early, Expected, and Very Late. There is an arrow labeled "Attentional Pulse" pointing to the Expected point on the graph.}]
Subsequent Experiments

Over a dozen experiments later
(e.g., Barnes & Jones, 2000; McAuley & Jones, 2002)

We Find:

• The profile vanishes with no context; sharpens with longer sequences.
• Holds for different rates.
• Reveals harmonic scaling of timing.
Different Induction Rates

Isochronous sequences

Three conditions: different induction rates.

<table>
<thead>
<tr>
<th>Condition</th>
<th>Induction IOIs</th>
<th>Standard durations</th>
</tr>
</thead>
<tbody>
<tr>
<td>1. 300 ms</td>
<td>300 ms</td>
<td>524, 600, 676 ms</td>
</tr>
<tr>
<td>2. 500 ms</td>
<td>500 ms</td>
<td>524, 600, 676 ms</td>
</tr>
<tr>
<td>3. 600 ms</td>
<td>600 ms</td>
<td>524, 600, 676 ms</td>
</tr>
</tbody>
</table>

This task required same vs different duration judgments
Conclude

*Back to Relative timing*

1. **Ratios** of induction IOIs to expected standards seem influential.

2. **Harmonic ratios** of 1:1 and 1:2 yield different profiles than the 500 ms with a more complex ratio.
Pitch Judgments

- We generalized this approach to attending to tone frequency.

- Does rhythmic context affect pitch judgments?
Task: Is a comparison pitch: “same”, “higher”, ”lower” than a standard pitch?

“Ignore distracting pitches!”

Standard pitch

Comparison pitch

Critical IOI

“higher”
“same”
“lower”

Independent Variable: Critical IOI
Rationale

- In a rhythmic context, attending is paced in time.

- Pitch judgments will be best for temporally expected tones: When critical IOI = context IOI.
Similar Predictions

A quadratic expectancy profile

Onset Time of Comparison

Very Early Expected Very Late

Attentional Pulse

PC
Experiment 1: Expectancy Profile

Onset Time of Comparison

Proportion Correct

P<.005
Related Experiments

Found a:

flattened expectancy profile
with irregular context timing
Recent findings

Suggest the importance of higher-order time structure ...

In focal attending over time levels
An Accent Rhythm

Higher time level: 4 IOIs

Accent Timing

Lower time level: IOI

Accented Tones
Accent Rhythm is Varied

Regular

Irregular

Accented Tones
Same Pitch Judgment Task

Critical IOI is Varied

Standard pitch

Comparison pitch
Two Oscillators?

*Remember those two circadian oscillators (Morning, Evening)?*

*Well, these are different!*

  *Shorter periods*
  *Sensitive to sound*
Two Different Oscillators

Two oscillators: Different periods

1. Lower-order oscillator, \( \text{period} \sim IOI \).
2. Accent oscillator, \( \text{period} \sim 4 IOIs \)

Which dominates?

If the Accent oscillator dominates, then only the Regular Rhythm will give an Expectancy Profile.
Two Oscillators

In listening to sound patterns

**Differential responses to sounds:**
Accented versus Unaccented

**Flexible attending:**
Dominance of one oscillator in focal attending.
Final Comments

Assumptions about rhythm and attentional synchrony can enhance theoretical power.

The dynamics of sound patterns together the dynamics of attending may explain real time tracking of events.

Entrainment to dynamic contexts may play a role in judgments of both time and pitch.
Two routes to Synchrony

1. **Anticipatory attending:**
   
   Attending heightens **prior** to an element.
   
   Oscillator **period** is important.

2. **Reactive attending:**
   
   If anticipation fails,
   
   Attending shifts **after** an element occurs.
   
   Oscillator **phase** is important.