Auditory Scene Analysis

Kimo Johnson
April 23, 2009
Auditory scene analysis

• Source segregation
  - Spatial separation
  - Spectral and temporal qualities

• Stream segregation
  - proximity: frequency or time
  - continuity: follow trajectory
  - similarity: frequency, timbre, intensity
  - symmetry and closure
Single sound organization

Figure by MIT OpenCourseWare.

adapted from Handel figure 7.2
Old-plus-new heuristic

Figures removed due to copyright restrictions.
Auditory stream

Figures removed due to copyright restrictions.
Stream segregation

- Proximity
- Continuity
- Similarity
- Symmetry and closure
Multiple sounds

adapted from Handel figure 7.4
Multiple sound organization

• Tone sequences
  - Vary parameters to cause perception of subsequences

• Conflicting organizations
  - Ambiguous sequences that put strategies in conflict
Proximity

Figure by MIT OpenCourseWare.
Results

from van Noorden, 1975
J.S. Bach

- Toccata and Fugue in D Minor ~1700
J.S. Bach

- Toccata and Fugue in D Minor ~1700

Figure by MIT OpenCourseWare.
Similarity

- Sounds are grouped by timbre
Similarity

- Sounds are grouped by timbre

Example

from *Music, Cognition, and Computerized Sound*, ed. Perry Cook
Competing organizations

Bregman and Pinker, 1978
Scale illusion

• Deutsch, 1975

Figure by MIT OpenCourseWare.
Scale illusion

• Deutsch, 1975

Demo

Figure by MIT OpenCourseWare.
Scale illusion

• Deutsch, 1975
Continuity

Kluender and Jenison 1992 - glides
Continuity

adapted from Bregman, 1990
Continuity

adapted from Bregman, 1990
Restoration

• Sasaki (1980) - familiar piano melodies
• Warren and Sherman (1974)
  - the *eel fell off the car
  - The *eal fell off the table
Frequency Graphs

Figures removed due to copyright restrictions.
Melodies

Diana Deutsch, 1972
Melodies

Diana Deutsch, 1972

melody 1
Melodies

Diana Deutsch, 1972

melody 1

melody 2

Tuesday, April 28, 2009
Music
Music

guitar and sax
Music and Speech Perception

Kimo Johnson
April 29, 2008
Linguistic universals

- Discreteness
- Semanticity
- Arbitrariness
- Openness
- Duality of patterning

Hockett, 1963
Music grammars

• Discreteness: $N$ pitches per octave

• Semanticity: scales, chords, keys

• Arbitrariness

• Openness

• Duality
Octave

- Frequency ratio 2:1
- Greatest number of identical overtones
- First overtone is 2 : 1
Sensory dissonance

Plomp and Levelt, 1965
Local consonance

Figure by MIT OpenCourseWare.

Sethares, 1993
Pythagorean comma

\[ f_1 = \frac{3}{2} f_0 \]

\[ f_i = \frac{3}{2} f_{i-1} \]
Pythagorean comma

\[
f_1 = \frac{3}{2} f_0
\]

\[
f_i = \frac{3}{2} f_{i-1}
\]

\[
\left(\frac{3}{2}\right)^{12} \approx 2^7
\]

\[
\frac{3^{12}}{2^{19}} \approx 1.0136
\]
# Pythagorean tuning

<table>
<thead>
<tr>
<th>C</th>
<th>C#</th>
<th>D</th>
<th>Eb</th>
<th>E</th>
<th>F</th>
<th>F#</th>
<th>G</th>
<th>G#</th>
<th>A</th>
<th>Bb</th>
<th>B</th>
</tr>
</thead>
<tbody>
<tr>
<td>1/256</td>
<td>9/8</td>
<td>32/27</td>
<td>81/64</td>
<td>4/3</td>
<td>729/512</td>
<td>3/2</td>
<td>128/81</td>
<td>27/16</td>
<td>16/9</td>
<td>243/128</td>
<td></td>
</tr>
</tbody>
</table>

\[
\begin{align*}
C : E &= \frac{81}{64} = 1.2656 \\
C : E &= \frac{5}{4} = 1.25 \\
C : Eb &= \frac{32}{27} = 1.1852 \\
C : Eb &= \frac{6}{5} = 1.20 \\
C# : F# &= 1.3515 \neq 1.333
\end{align*}
\]
## Pythagorean tuning

<table>
<thead>
<tr>
<th></th>
<th>C</th>
<th>C#</th>
<th>D</th>
<th>Eb</th>
<th>E</th>
<th>F</th>
<th>F#</th>
<th>G</th>
<th>G#</th>
<th>A</th>
<th>Bb</th>
<th>B</th>
</tr>
</thead>
<tbody>
<tr>
<td>1/1</td>
<td>256/243</td>
<td>9/8</td>
<td>32/27</td>
<td>81/64</td>
<td>4/3</td>
<td>729/512</td>
<td>3/2</td>
<td>128/81</td>
<td>27/16</td>
<td>16/9</td>
<td>243/128</td>
<td></td>
</tr>
</tbody>
</table>

- C : E = \(\frac{81}{64}\) = 1.2656
- C : E = \(\frac{5}{4}\) = 1.25

- C : Eb = \(\frac{32}{27}\) = 1.1852
- C : Eb = \(\frac{6}{5}\) = 1.20

- C# : F# = 1.3515 ≠ 1.333

*Wolf*
Other tuning systems

• Just diatonic

• Meantone (1400)

• Well-temperaments
  - Werckmeister (1645-1706)
  - Young (1773-1829)

• Equal temperament (~1900)
Optimal well-temperament

Polansky et. al, 2008