

MIT OpenCourseWare  
<http://ocw.mit.edu>

9.35 Sensation And Perception  
Spring 2009

For information about citing these materials or our Terms of Use, visit: <http://ocw.mit.edu/terms>.

# 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

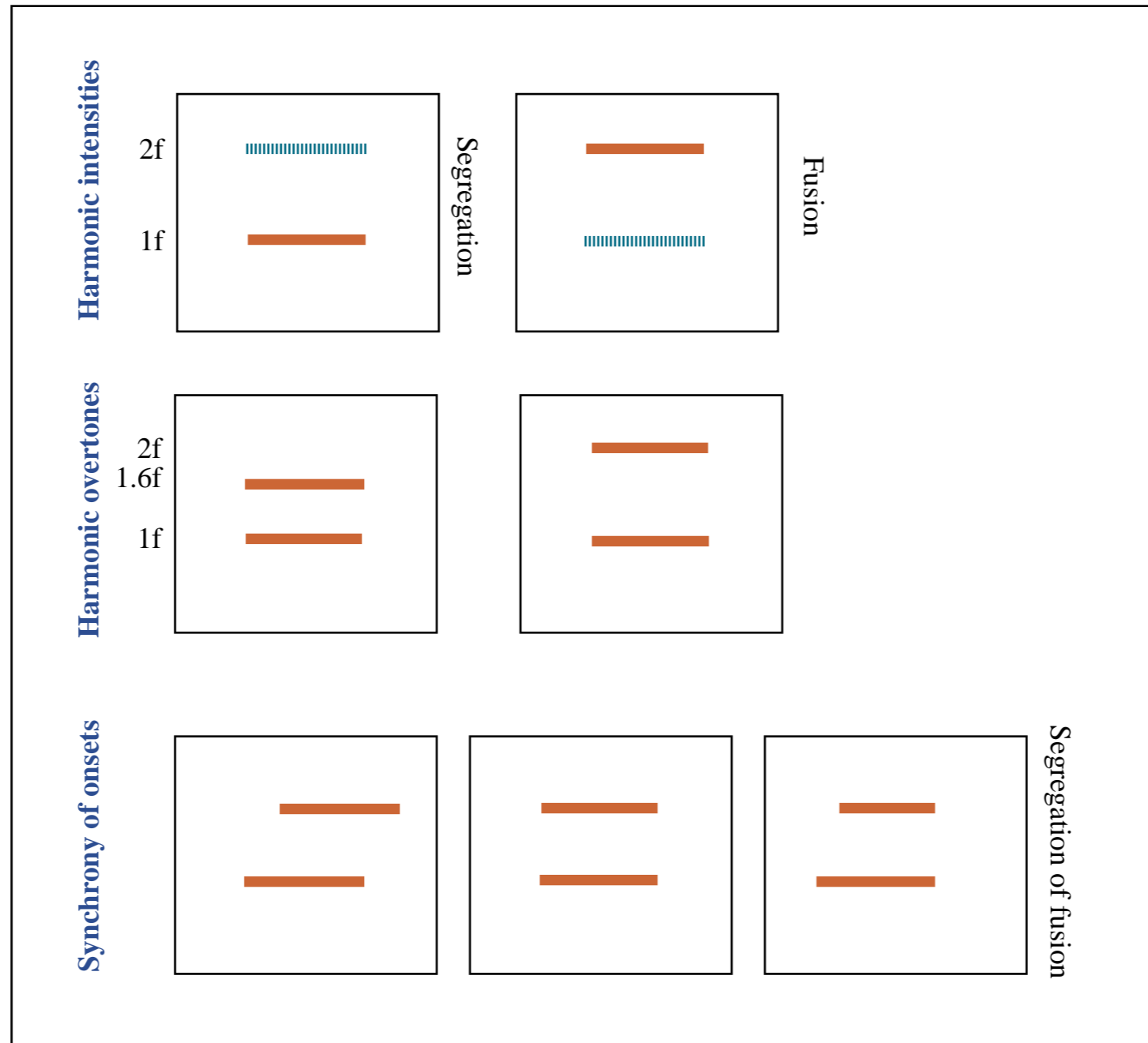


Figure by MIT OpenCourseWare.

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

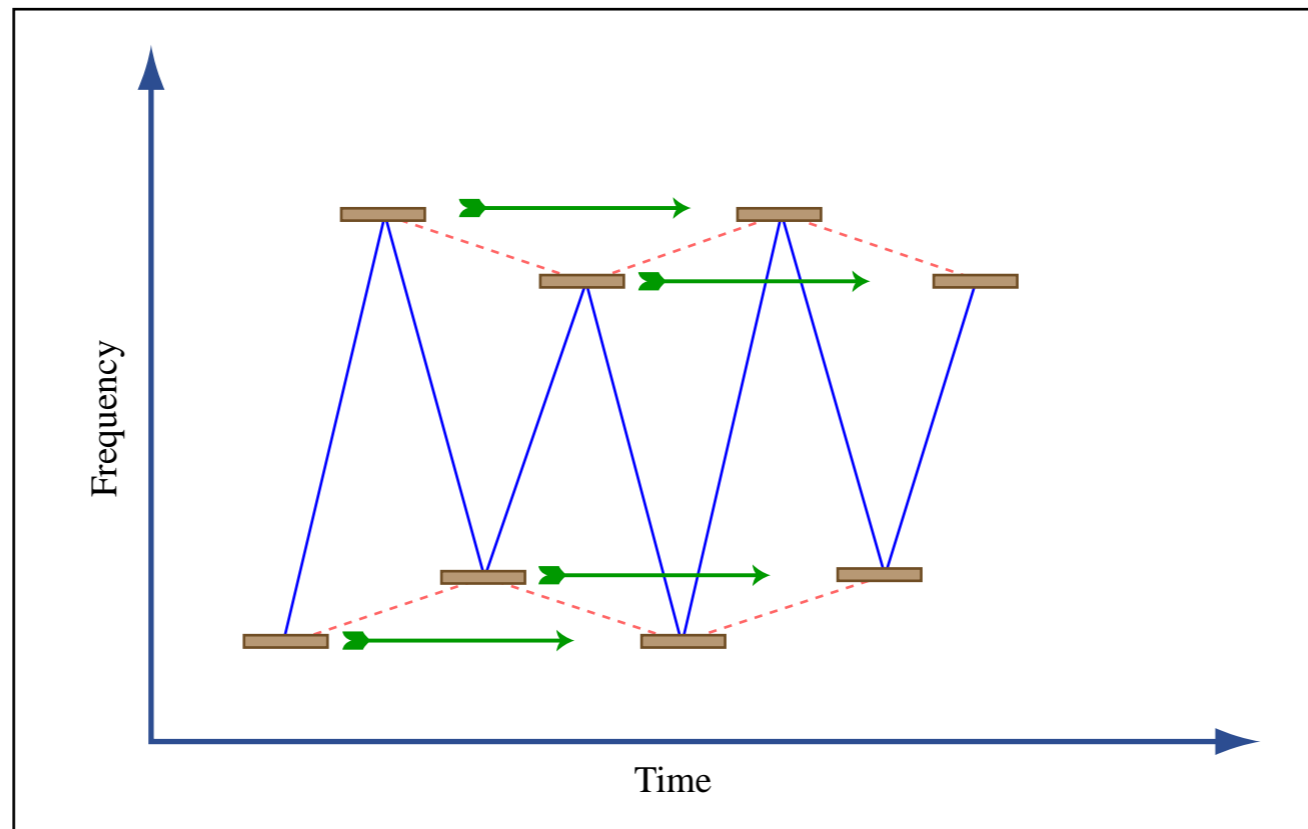


Figure by MIT OpenCourseWare.

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

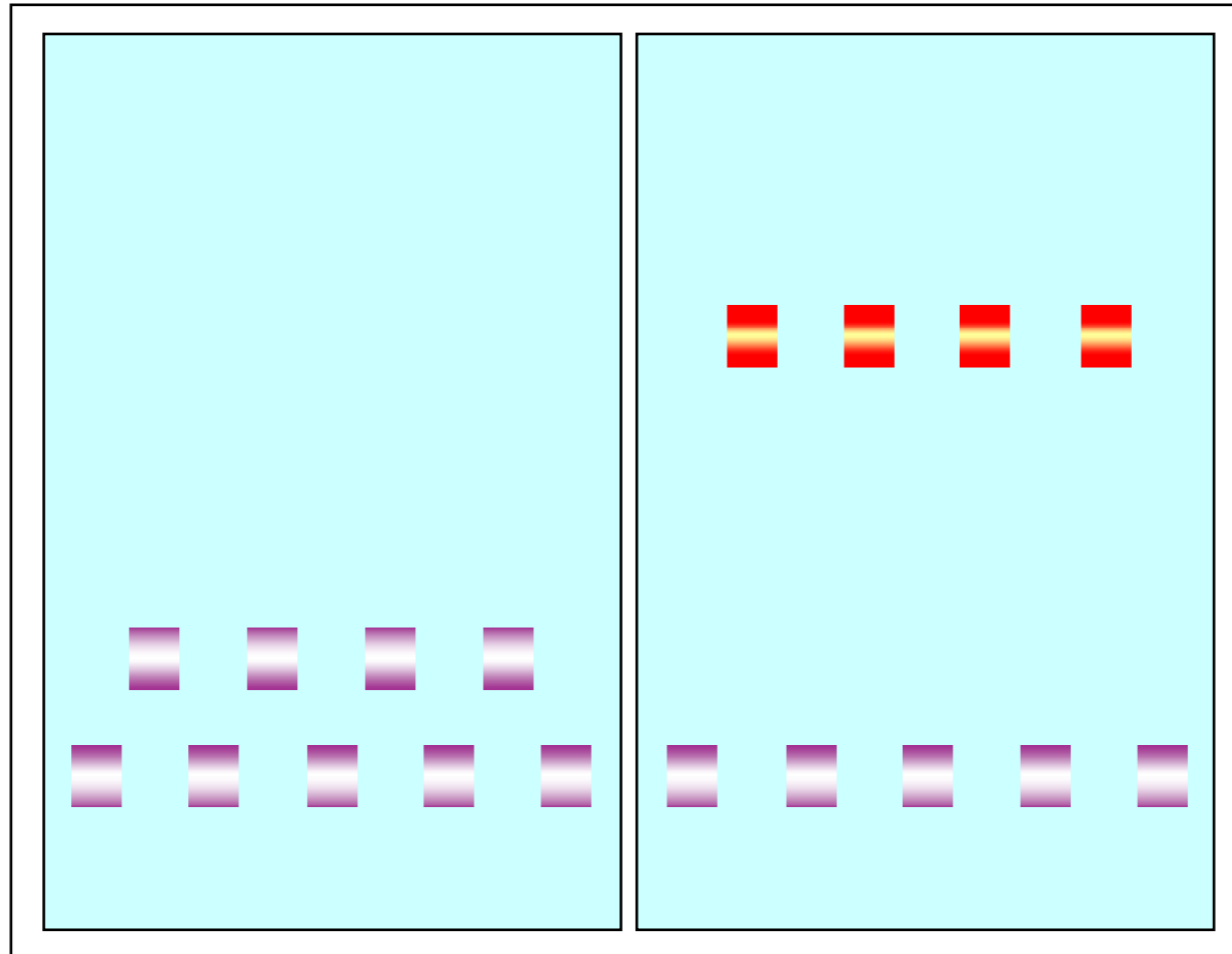


Figure by MIT OpenCourseWare.

# Results

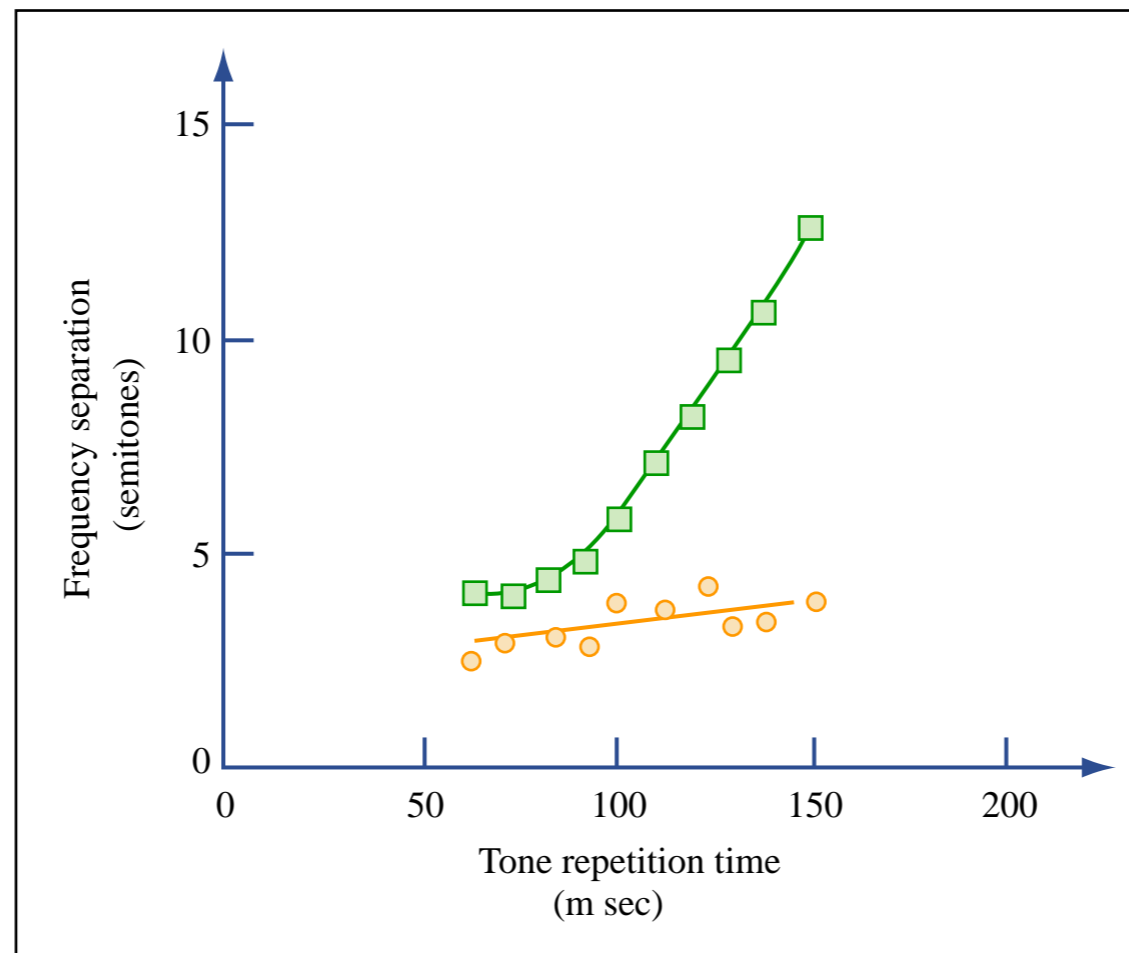


Figure by MIT OpenCourseWare.

from van Noorden, 1975

# J.S. Bach

- Toccata and Fugue in D Minor ~1700

# J.S. Bach

- Toccata and Fugue in D Minor ~1700

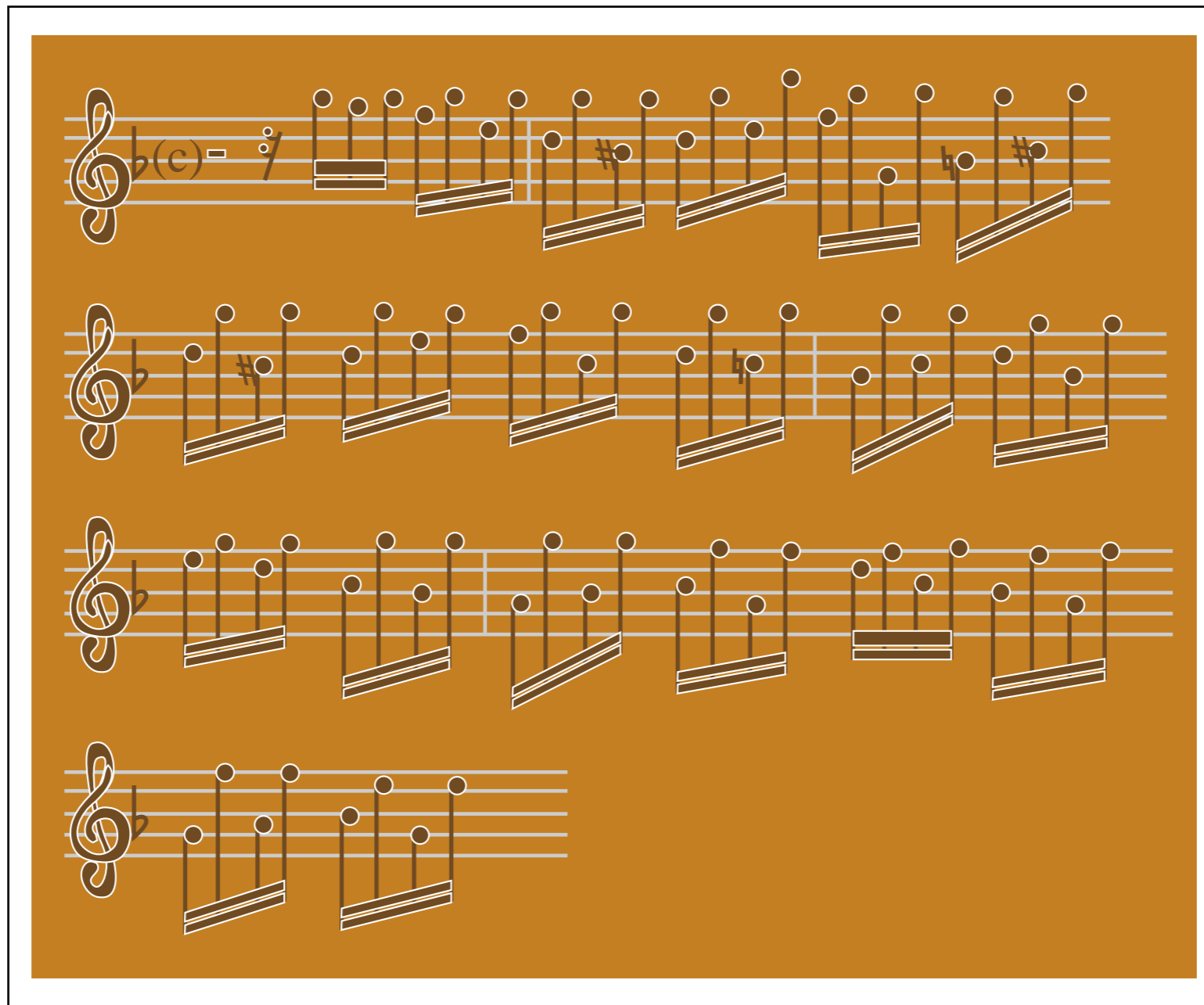


Figure by MIT OpenCourseWare.

# Similarity

- Sounds are grouped by timbre

# Similarity

- Sounds are grouped by timbre

## Example

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

# Competing organizations

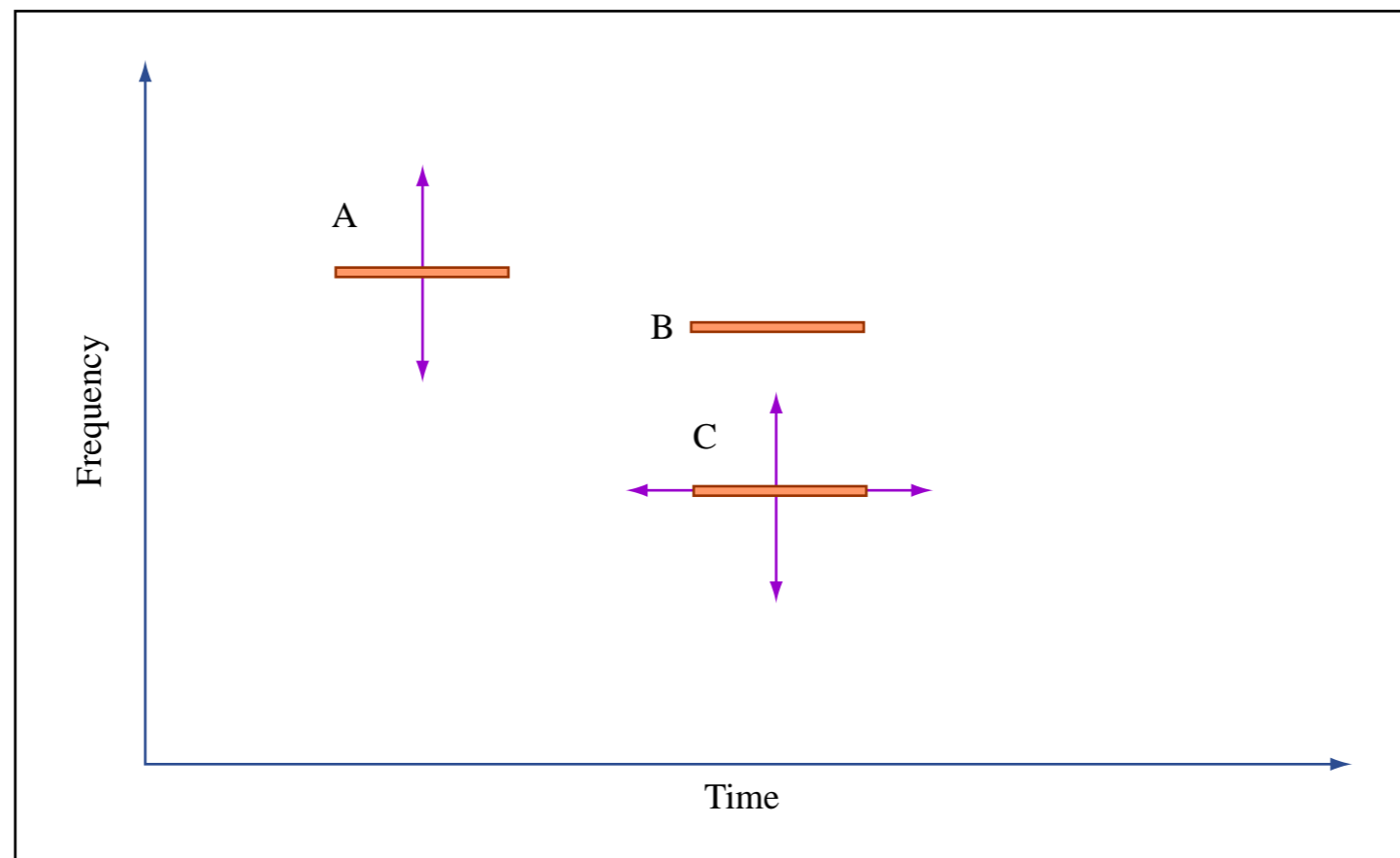


Figure by MIT OpenCourseWare.

Bregman and Pinker, 1978



# Scale illusion

- Deutsch, 1975

The image displays two parts of a musical scale illusion on a brown background. **Part 1** consists of two staves of musical notation. The top staff has notes with 'L' and 'R' labels above them, indicating the hand used to play each note. The bottom staff has notes with 'R' and 'L' labels below them. **Part 2** consists of two staves of musical notation, each starting with a treble clef. The notes in Part 2 are arranged to create an auditory illusion of a scale.

Figure by MIT OpenCourseWare.

# Scale illusion

- Deutsch, 1975

The image displays two parts of a musical exercise on a brown background. **Part 1** consists of two staves of music. The first staff has notes with fingerings: L, R, L, R, L, R, L, R, L, R, L, R, L, R. The second staff has fingerings: R, L, R, L, R, L, R, L, R, L, R, L, R, L. **Part 2** consists of two staves of music. The first staff shows a standard scale with notes on the treble clef. The second staff shows a standard scale with notes on the bass clef.

Figure by MIT OpenCourseWare.

## Demo

# Scale illusion

- Deutsch, 1975

The image shows two parts of a musical scale illusion. Part 1 consists of two staves of musical notation. The top staff has notes with fingerings: L, R, L, R, L, R, L, R, L, R, L, R, L, R. The bottom staff has notes with fingerings: R, L, R, L, R, L, R, L, R, L, R, L, R, L. Part 2 consists of two staves of musical notation showing the same sequence of notes without fingerings.

Figure by MIT OpenCourseWare.

The image shows Part 3 and a Demo. Part 3 consists of two staves of musical notation showing the same sequence of notes without fingerings. The Demo consists of two staves of musical notation showing the same sequence of notes without fingerings.

Figure by MIT OpenCourseWare.

# Continuity

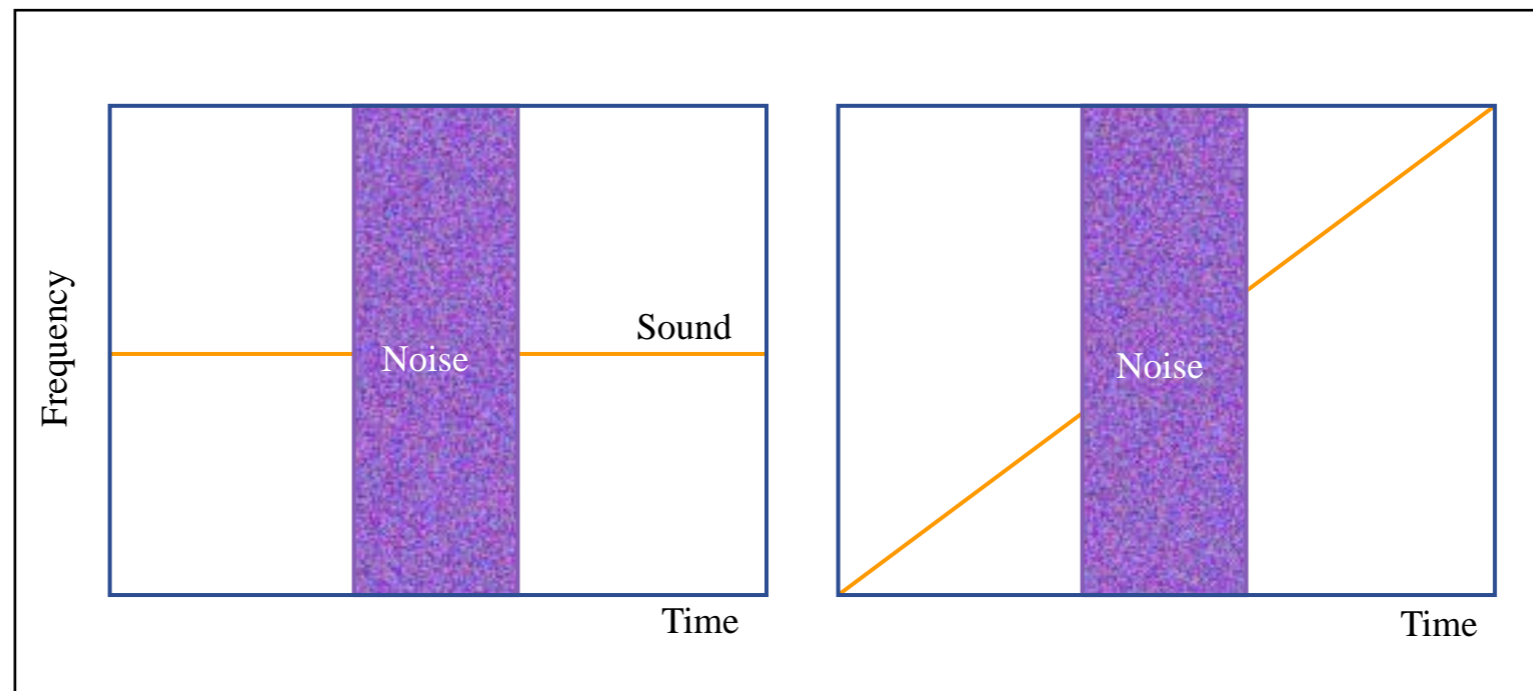
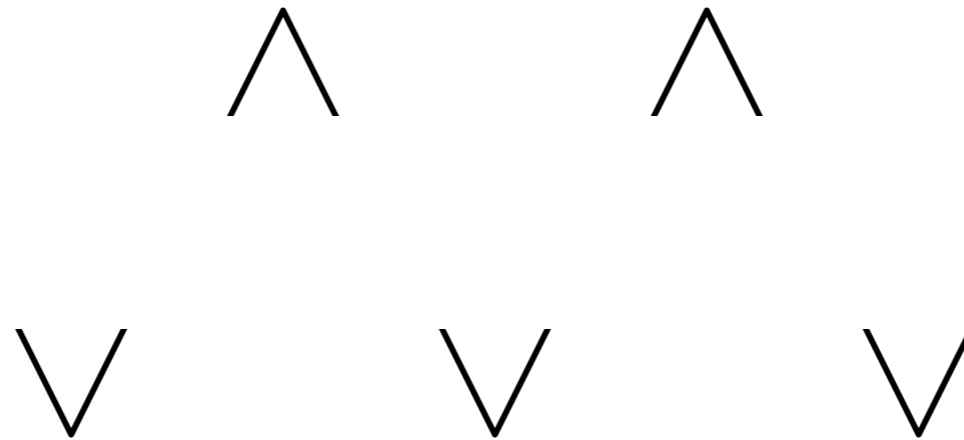


Figure by MIT OpenCourseWare.

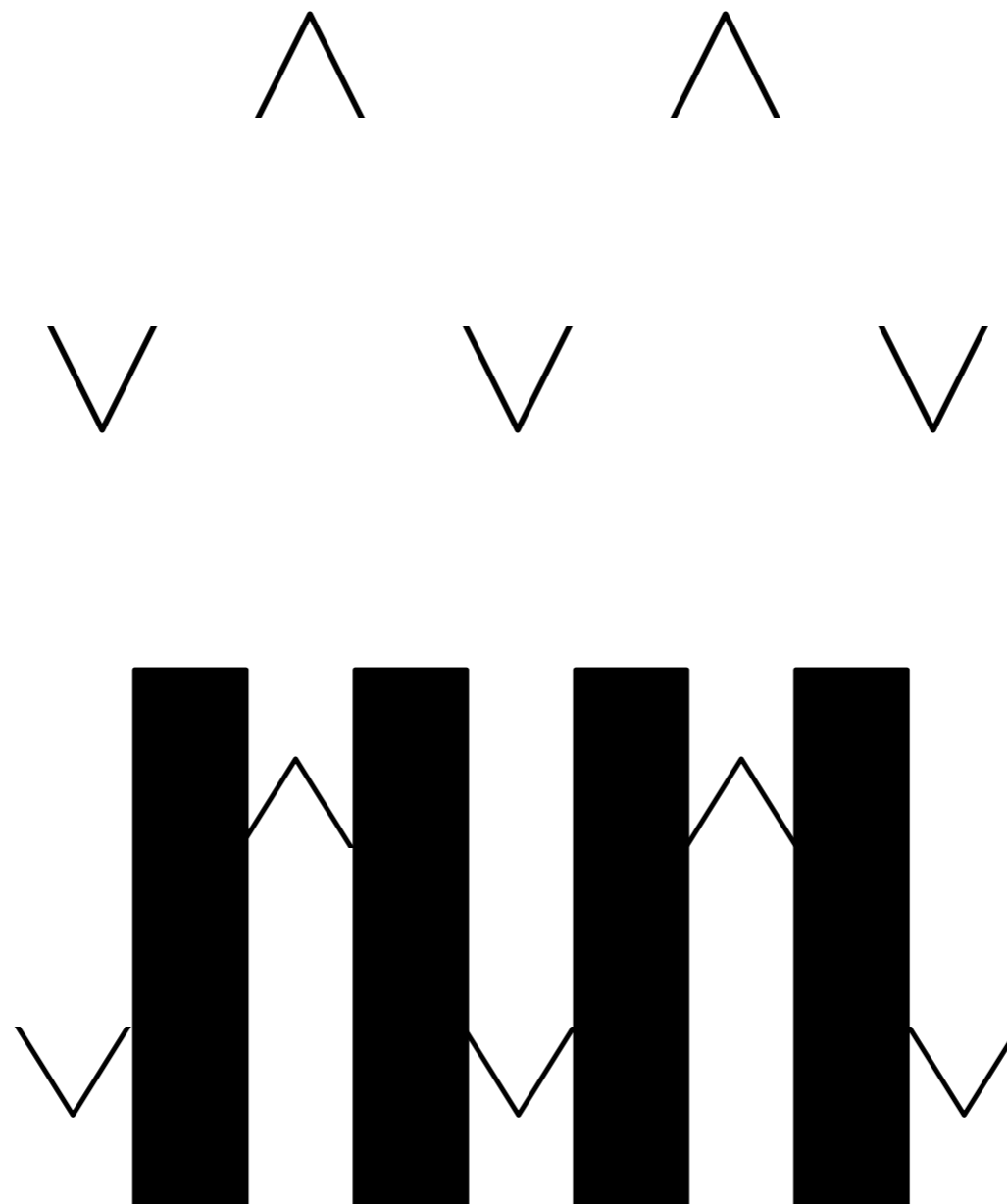
Kluender and Jenison 1992 - glides

# 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 I

# Melodies

Diana Deutsch, 1972

melody 1

melody 2

# 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

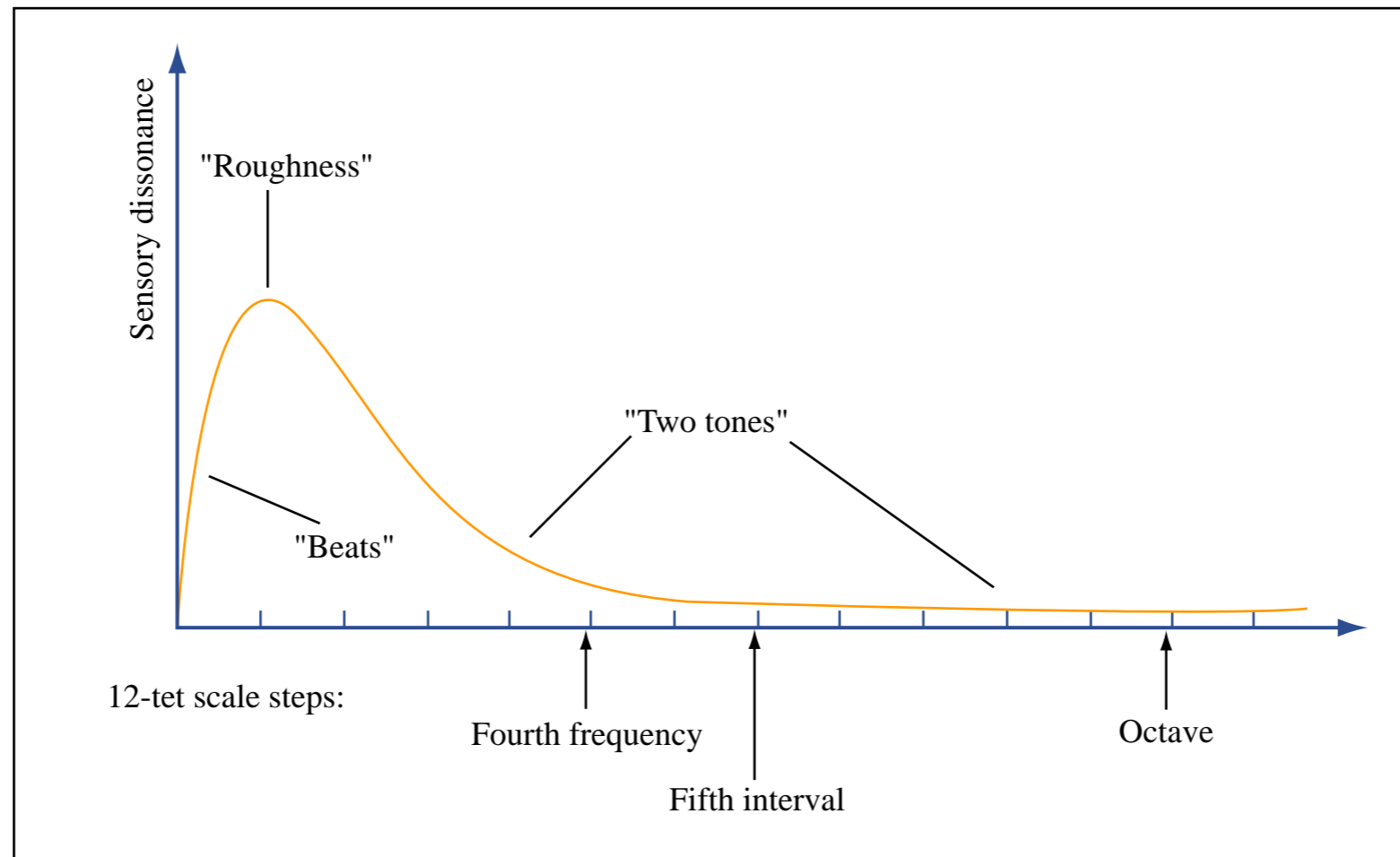


Figure by MIT OpenCourseWare.

## Plomp and Levelt, 1965

# Local consonance

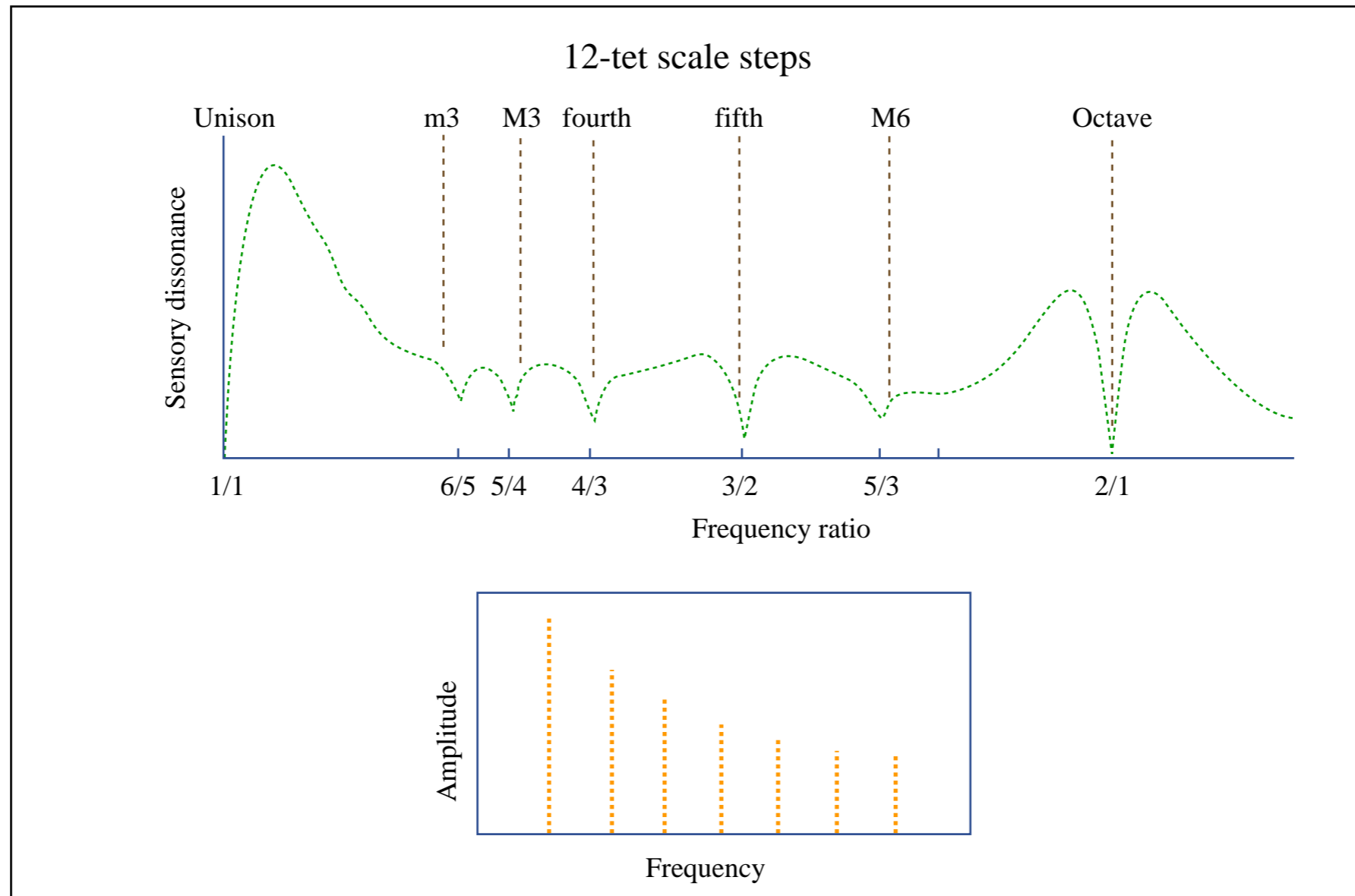
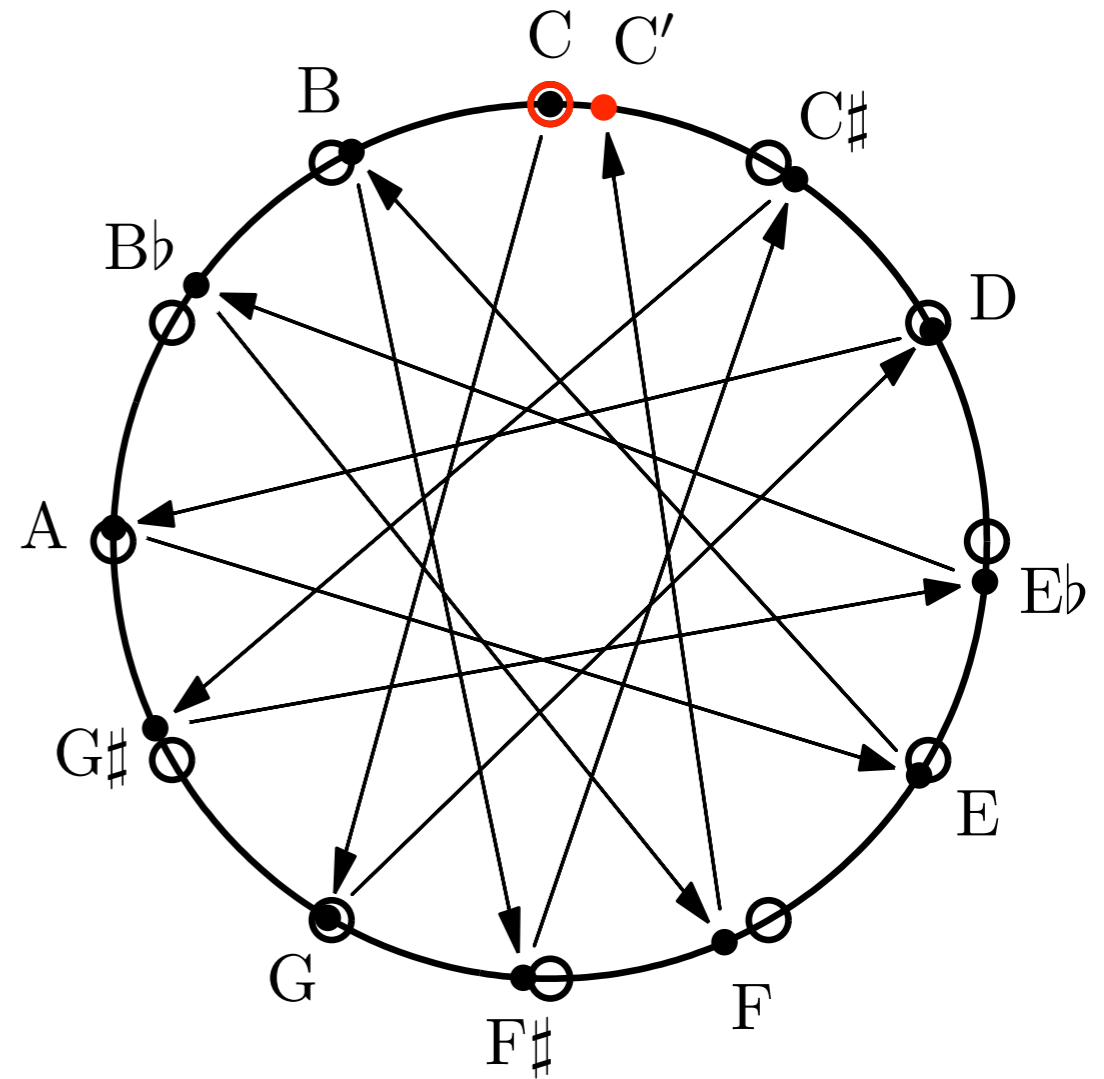


Figure by MIT OpenCourseWare.

Sethares, 1993

# Pythagorean comma

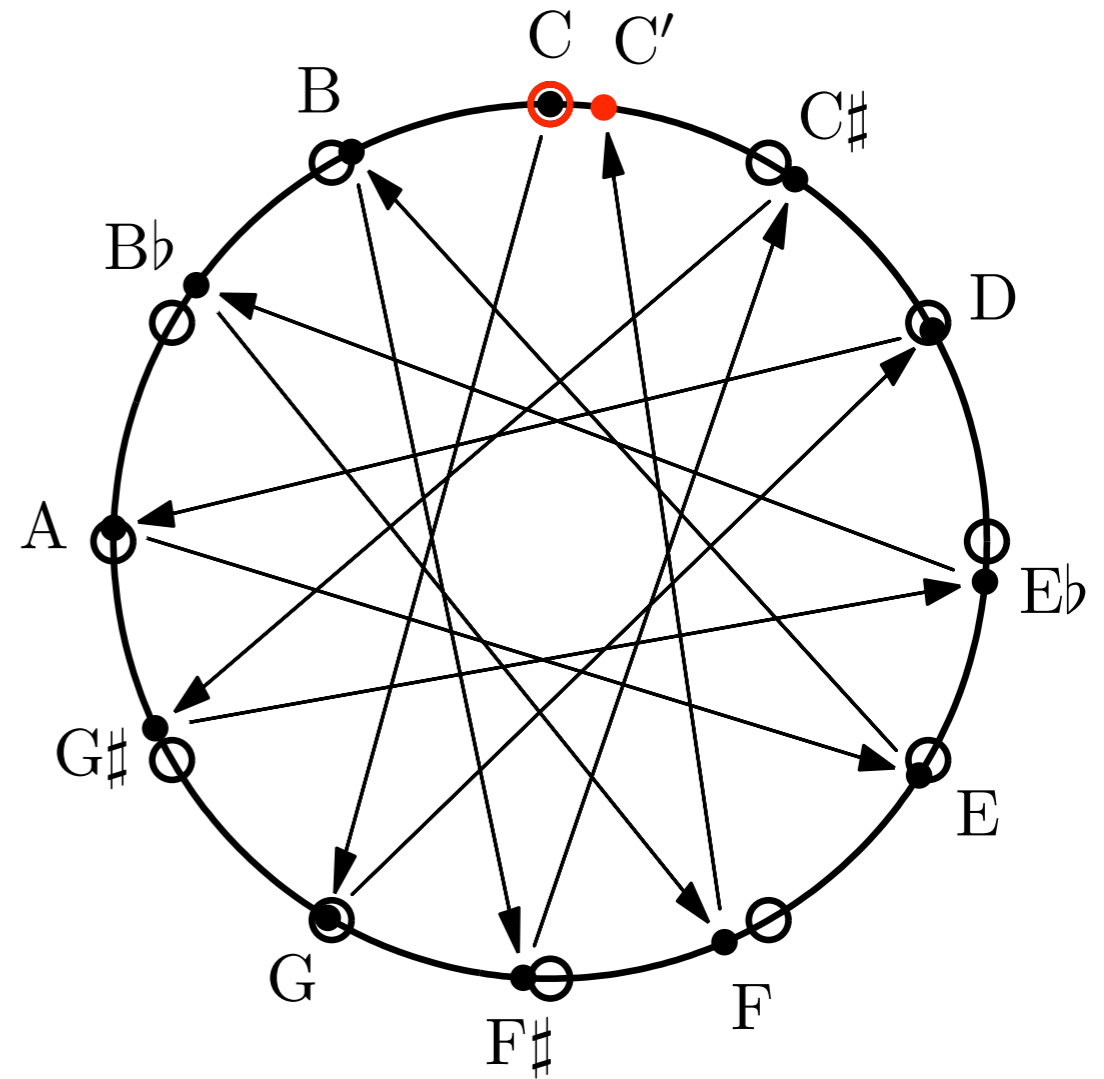
$$f_1 = \frac{3}{2}f_0$$
$$f_i = \frac{3}{2}f_{i-1}$$



# 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

C	C#	D	E $\flat$	E	F	F#	G	G#	A	B $\flat$	B
I/	256/	9/	32/	81/	4/	729/	3/	128/	27/	16/	243/
I	243	8	27	64	3	512	2	81	16	9	128

$$C : E = 81/64 = 1.2656$$

$$C : E = 5/4 = 1.25$$

$$C : E\flat = 32/27 = 1.1852$$

$$C : E\flat = 6/5 = 1.20$$

$$C\# : F\# = 1.3515 \neq 1.333$$

# Pythagorean tuning

C	C#	D	E <sub>b</sub>	E	F	F#	G	G#	A	B <sub>b</sub>	B
I/	256/	9/	32/	81/	4/	729/	3/	128/	27/	16/	243/
I	243	8	27	64	3	512	2	81	16	9	128

$$C : E = 81/64 = 1.2656$$

$$C : E = 5/4 = 1.25$$

$$C : E_b = 32/27 = 1.1852$$

$$C : E_b = 6/5 = 1.20$$

$$C\# : F\# = 1.3515 \neq 1.333$$

wolf

# Other tuning systems

- Just diatonic
- Meantone (1400)
- Well-temperaments
  - Werckmeister (1645-1706)
  - Young (1773-1829)
- Equal temperament (~1900)



# Optimal well-temperament

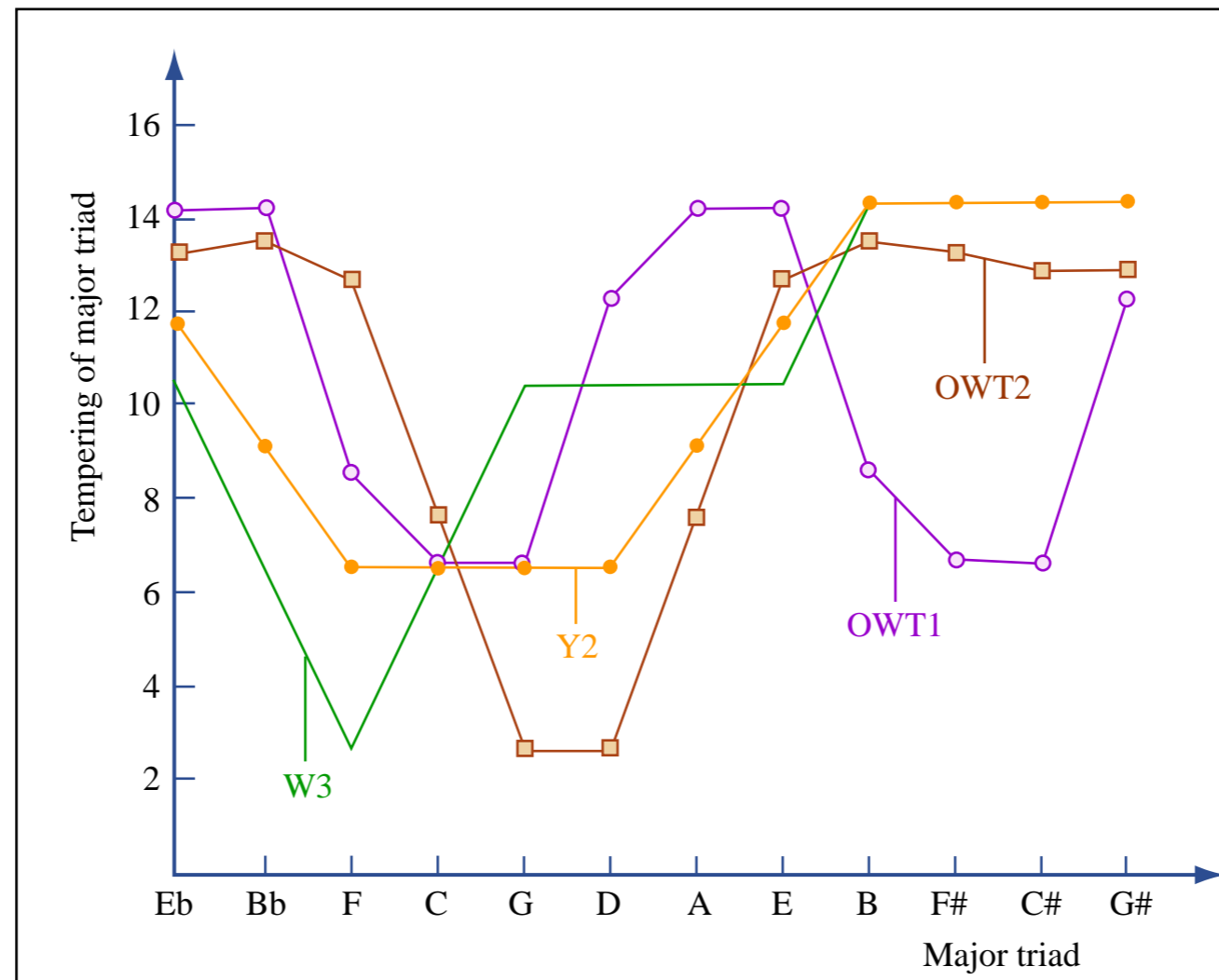


Figure by MIT OpenCourseWare.

Polansky et. al, 2008