



Classification Tasks

What can we classify?

What can we identify from music?



Classification Tasks

Genre & Emotion

Genre Classification

Where does one begin?



What does “genre” (musically) mean?

How do you describe “genre”?

Genre Classification

What assumptions do we make?

Classes are **acoustically** and **perceptually** separable

How do we go about selecting relevant acoustic features and parameters thereof?

Genre Classification

Assumption: **Classes are separable**

Perceptual



Genre 1

Genre 2

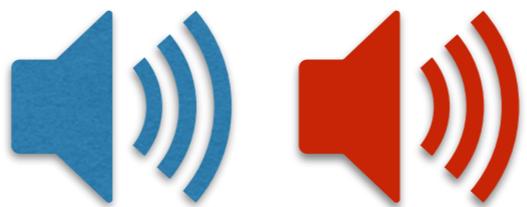


what are the perceptual features that distinguish these genres?

Genre Classification

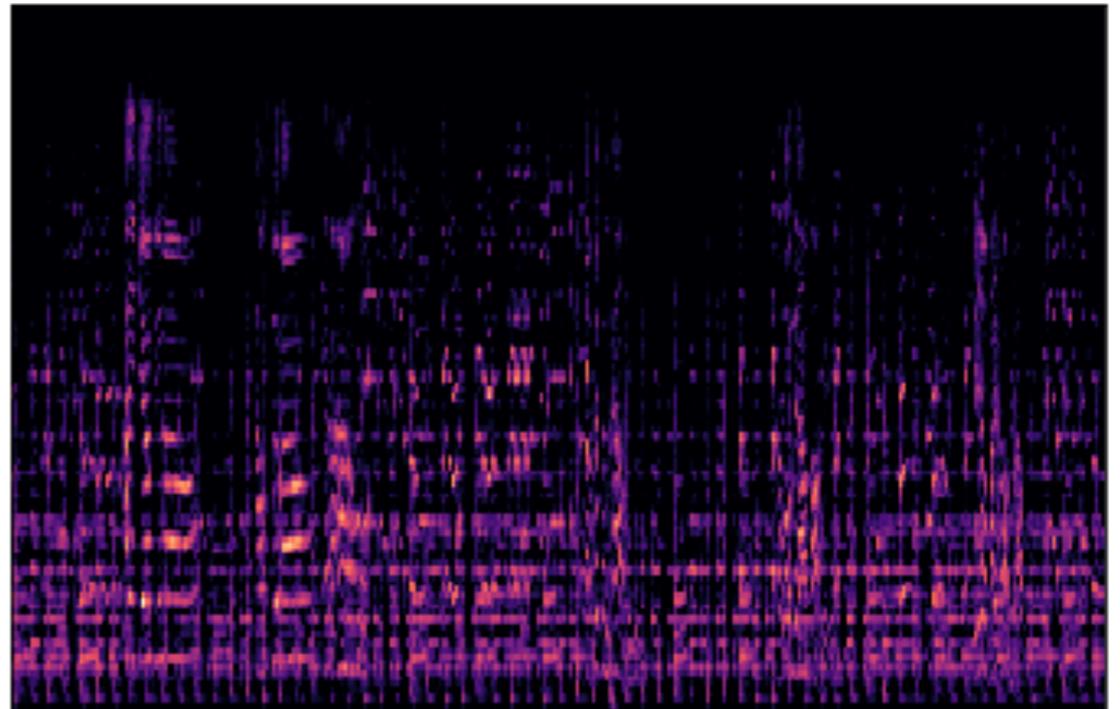
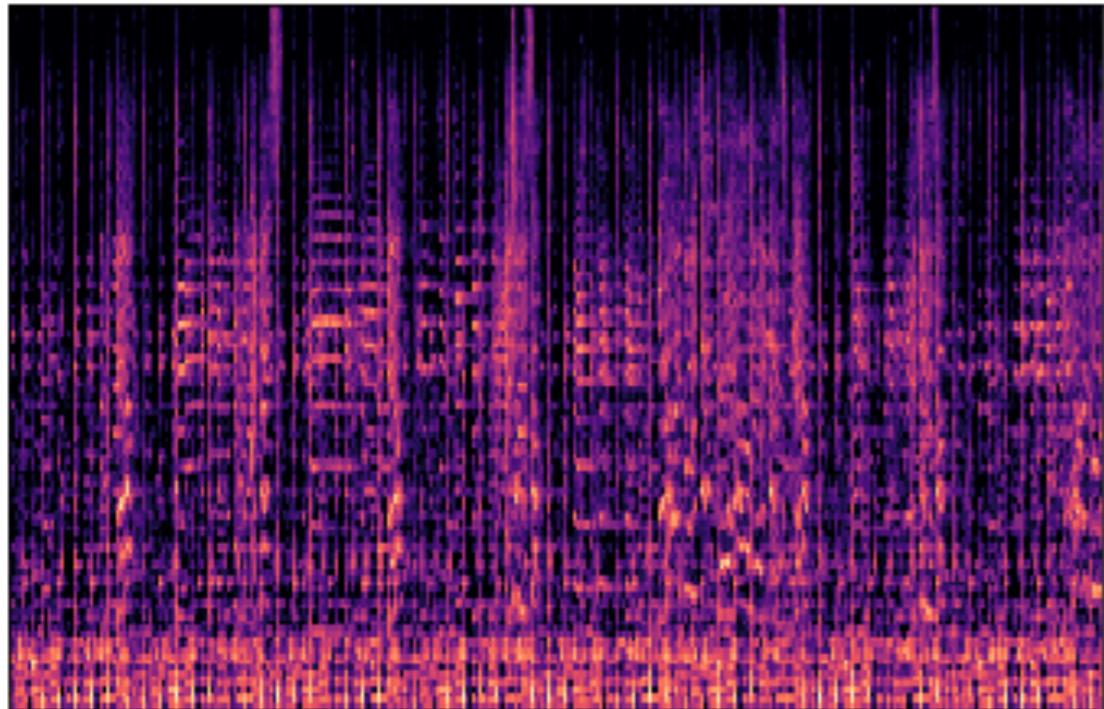
Assumption: **Classes are separable**

Genre 1



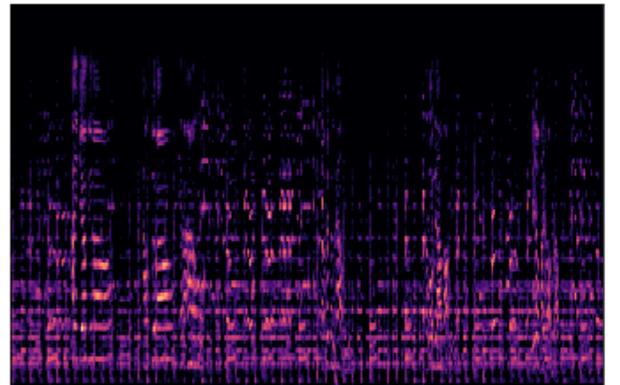
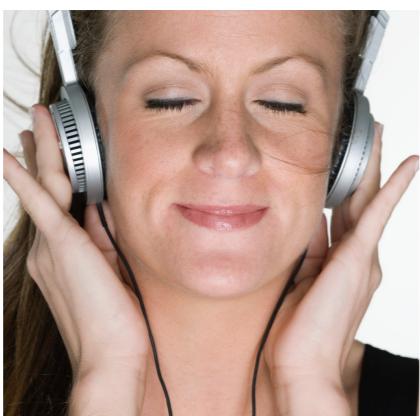
Genre 2

Acoustically

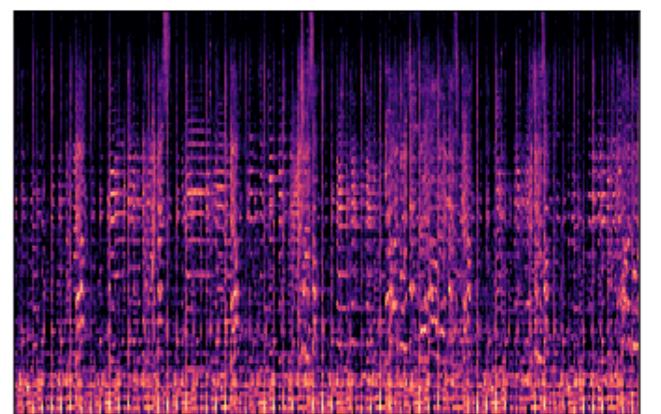


Genre Classification

based on perceptual features can you identify the acoustic features ?



Genre 1



Genre 2

Genre Classification

So what sort of features do i choose?

Timbre

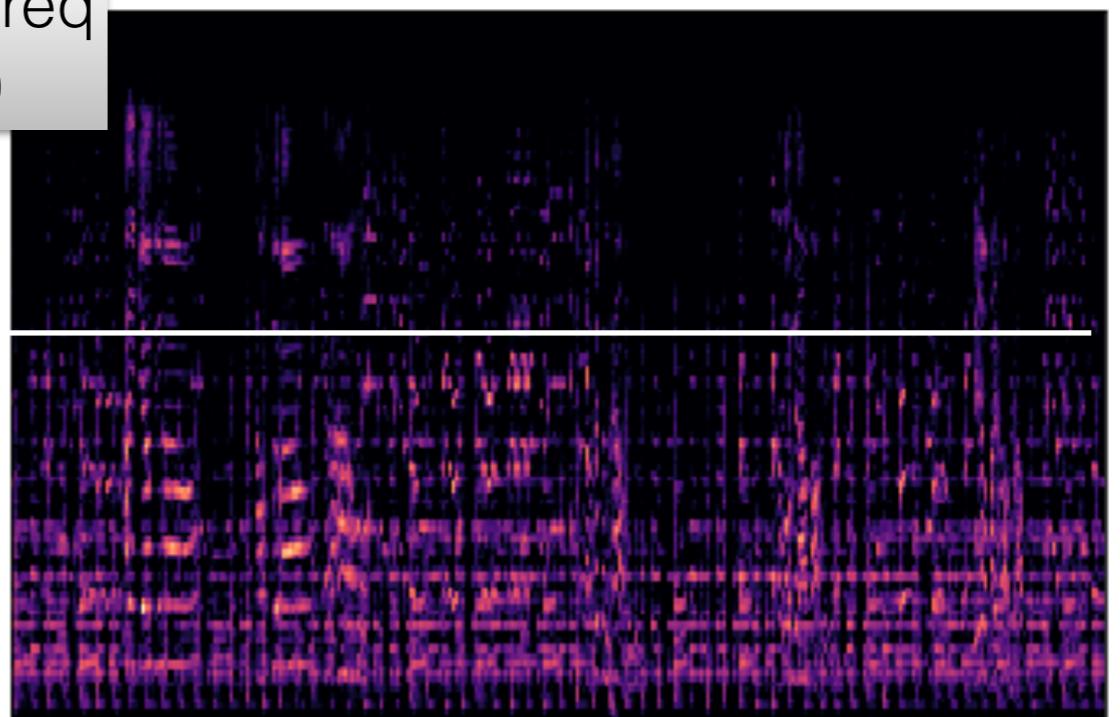
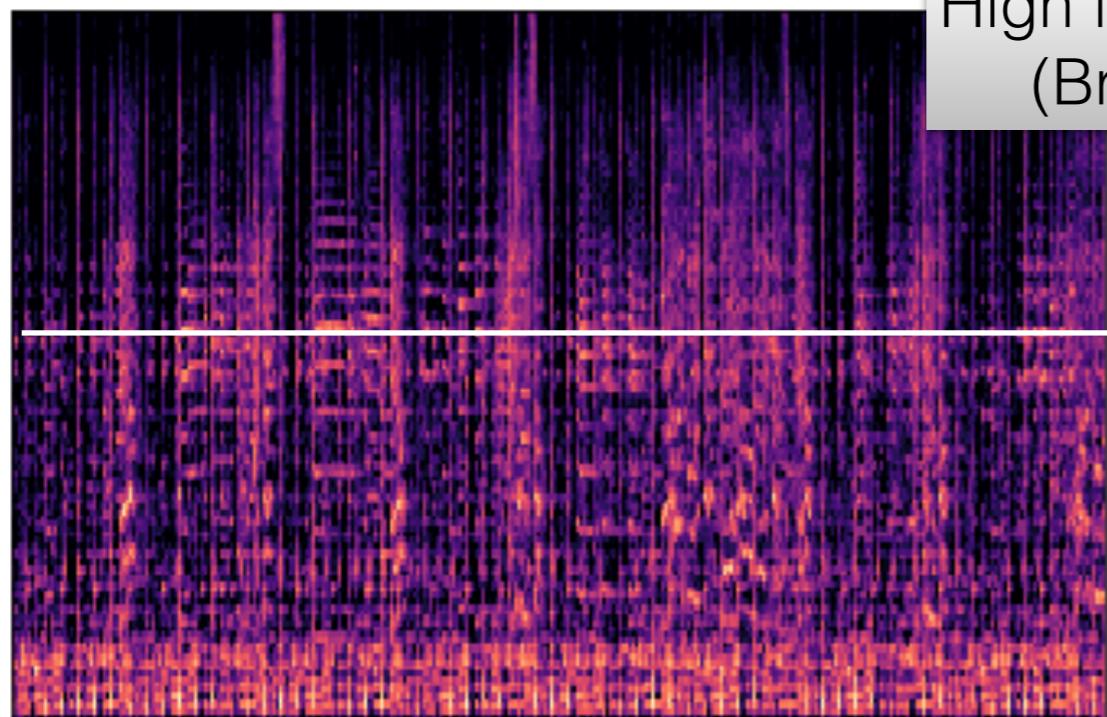
Rhythm

Key/Tonality

Genre 2

Genre 1

High freq/Low freq
(Brightness)



Spectral Centroid

Genre Classification

So what sort of features do i choose?

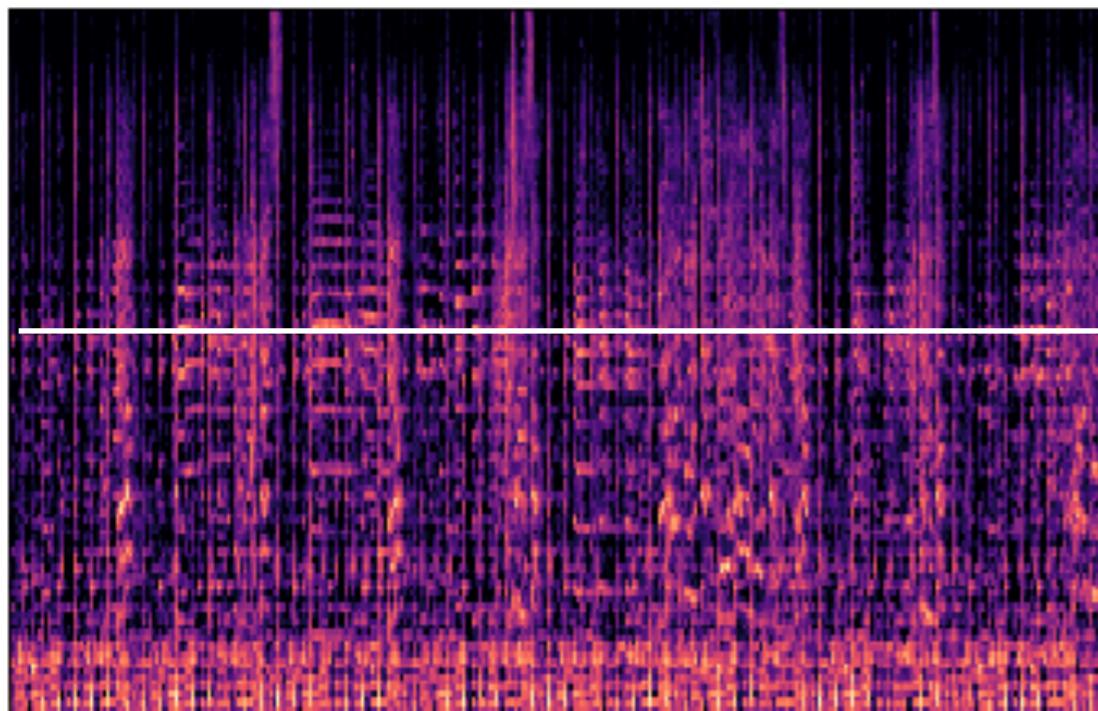
Timbre

Rhythm

Key/Tonality

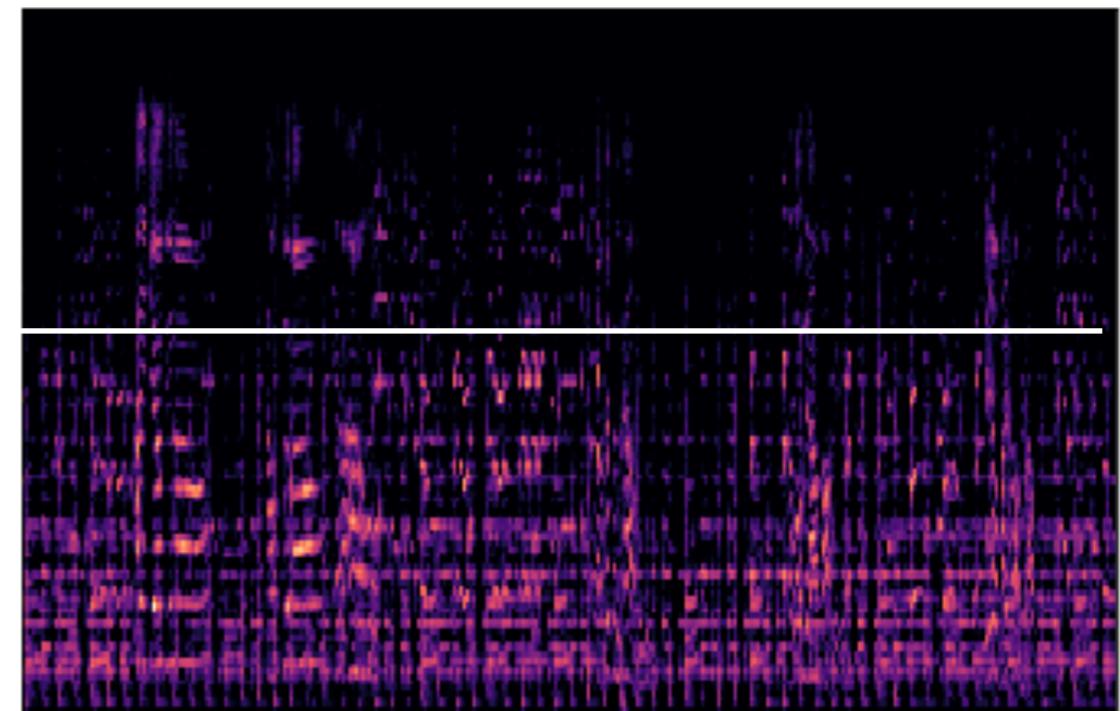
Genre 2

Genre 1



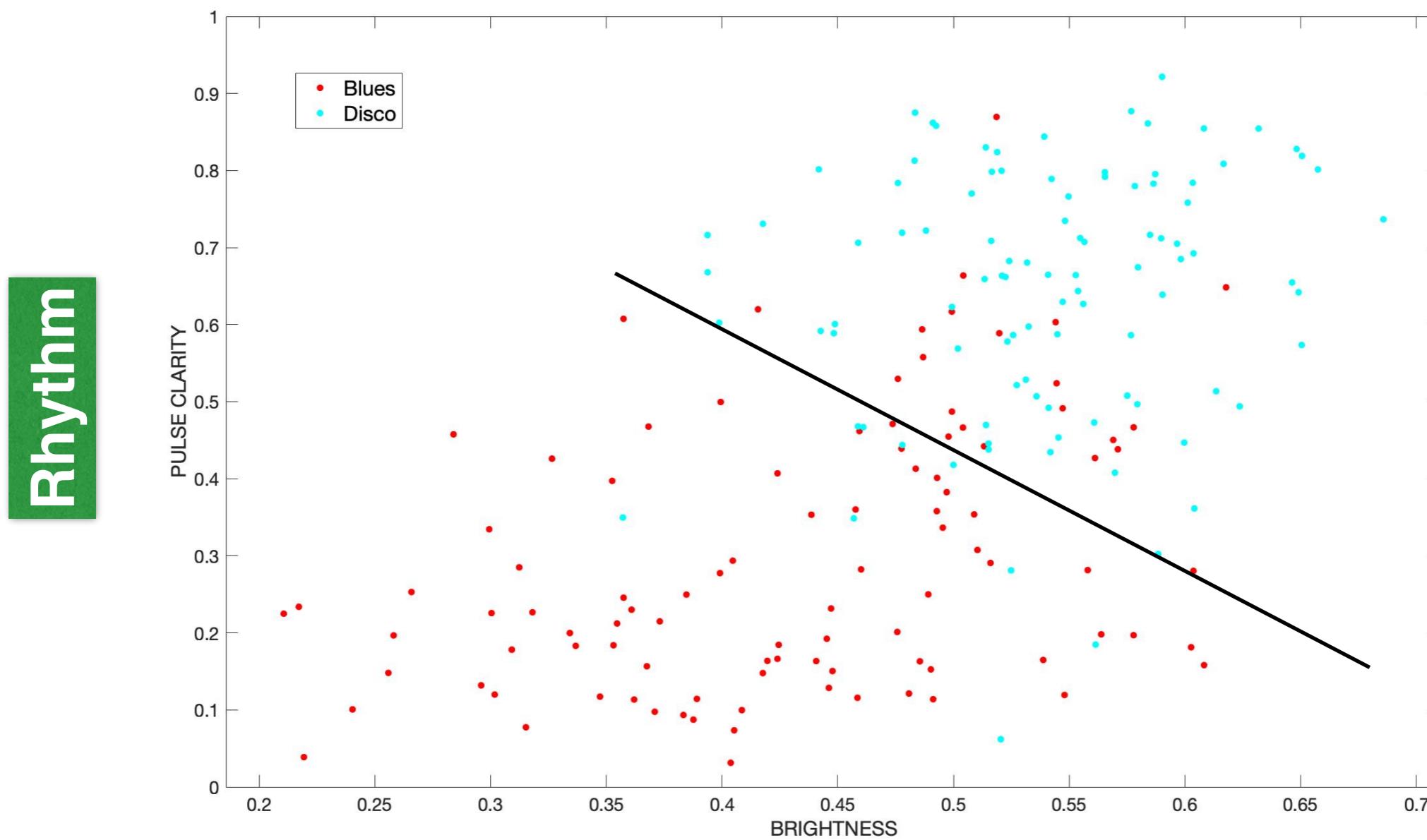
Tempo?

Pulse Clarity?



Genre Classification

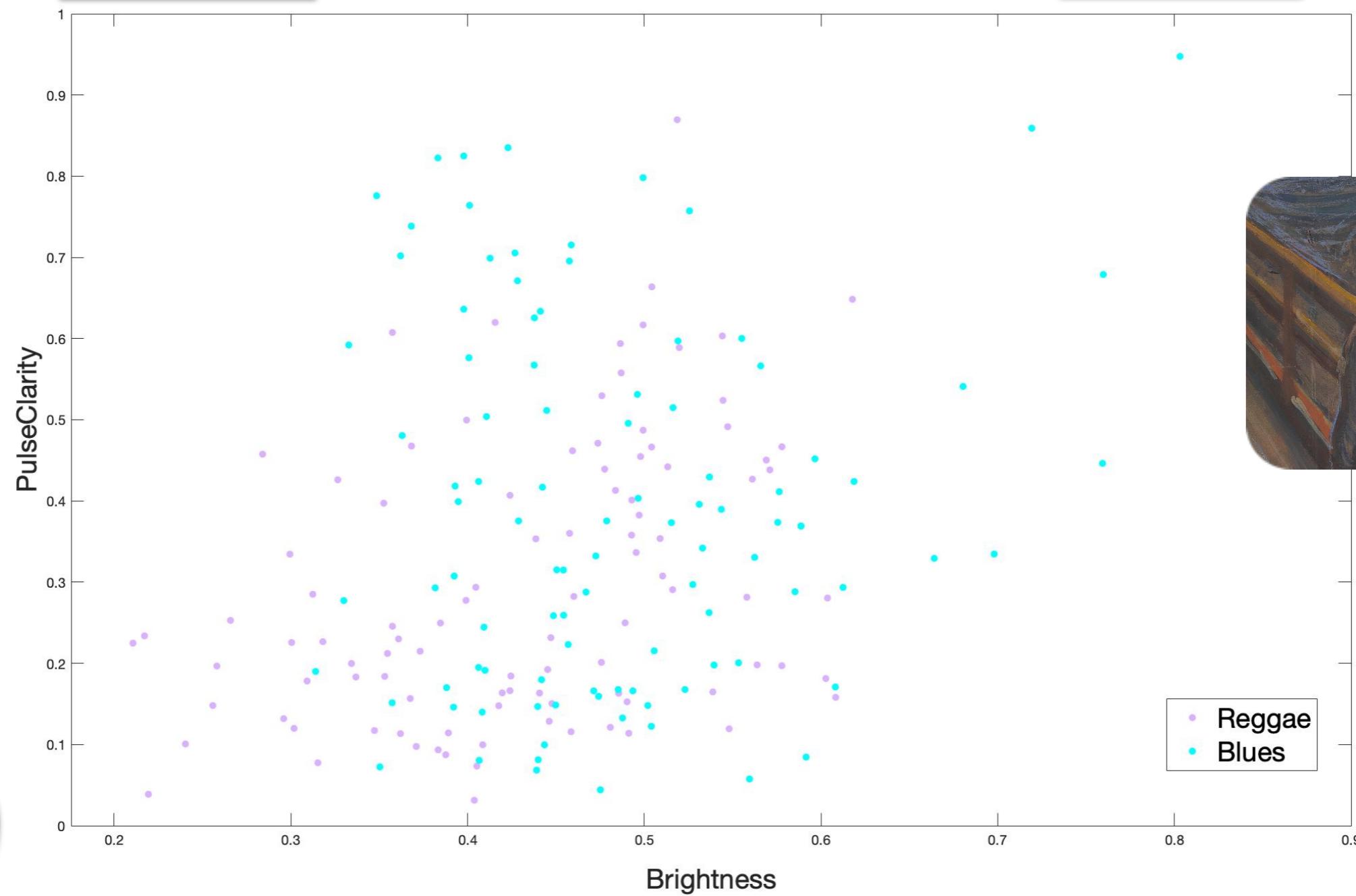
Acoustic feature selection



Timbre

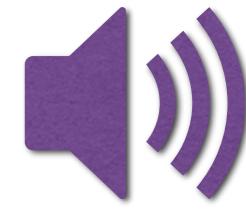
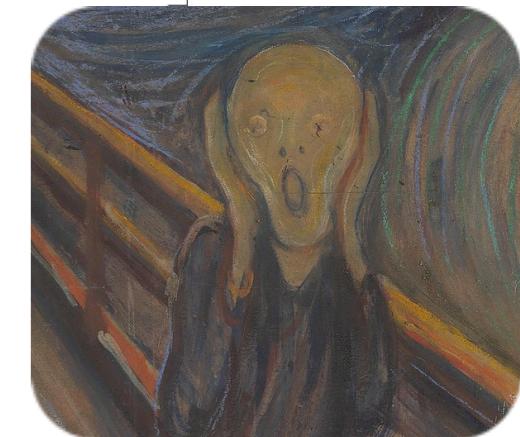
Genre Classification

Genre 1 Acoustic feature selection **Genre 2**



Timbre

Rhythm

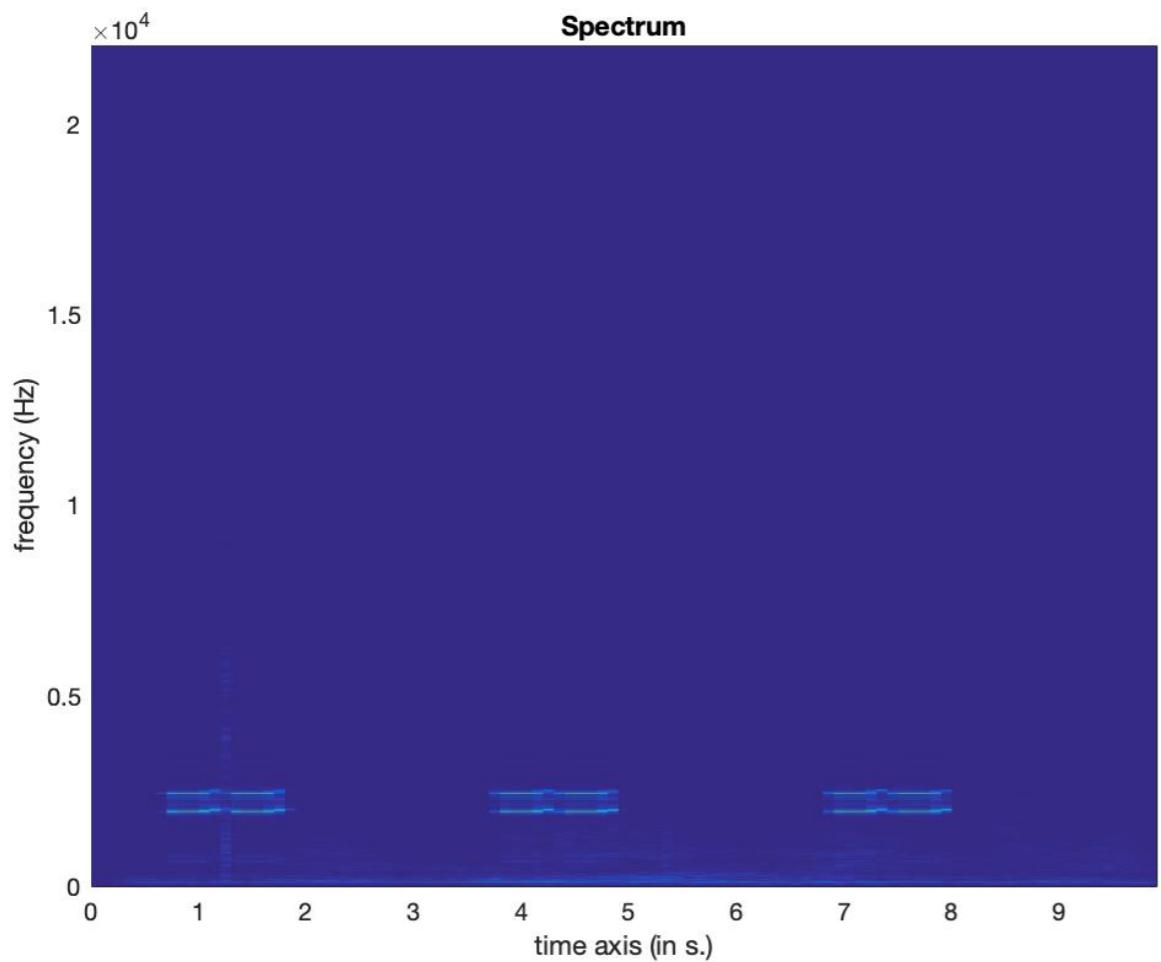




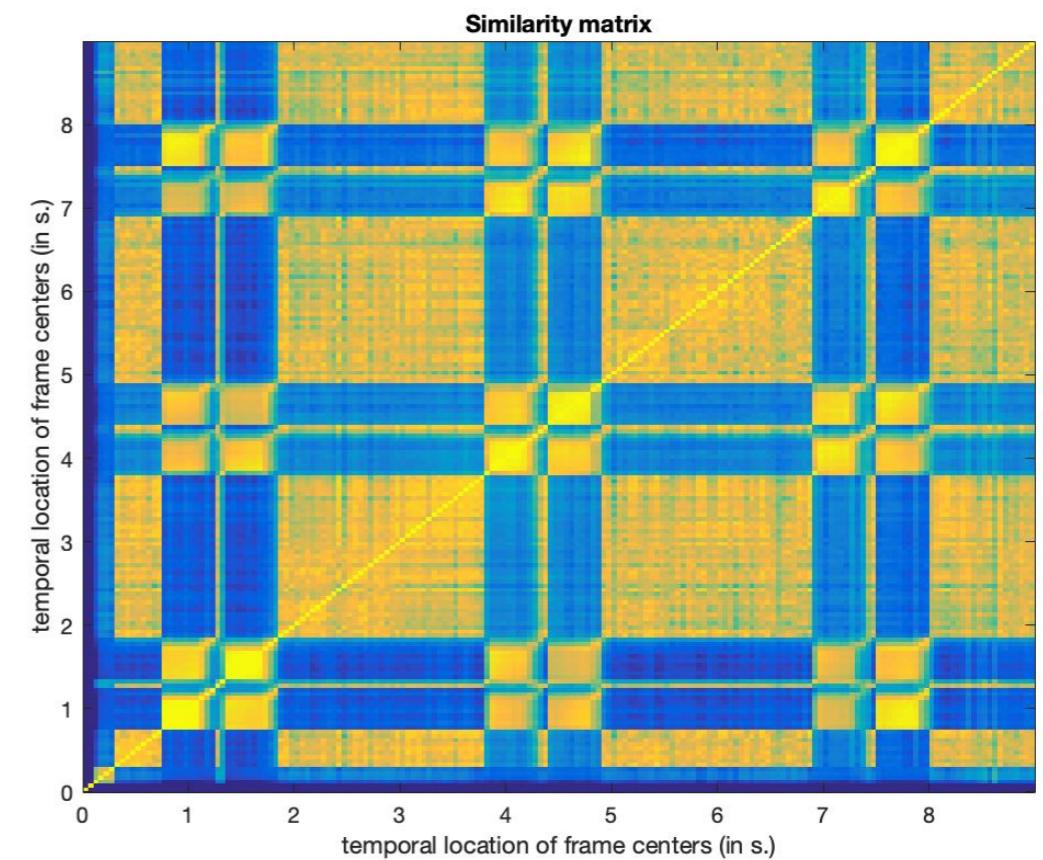
Visualizing Genres (Structure) guess the genre?



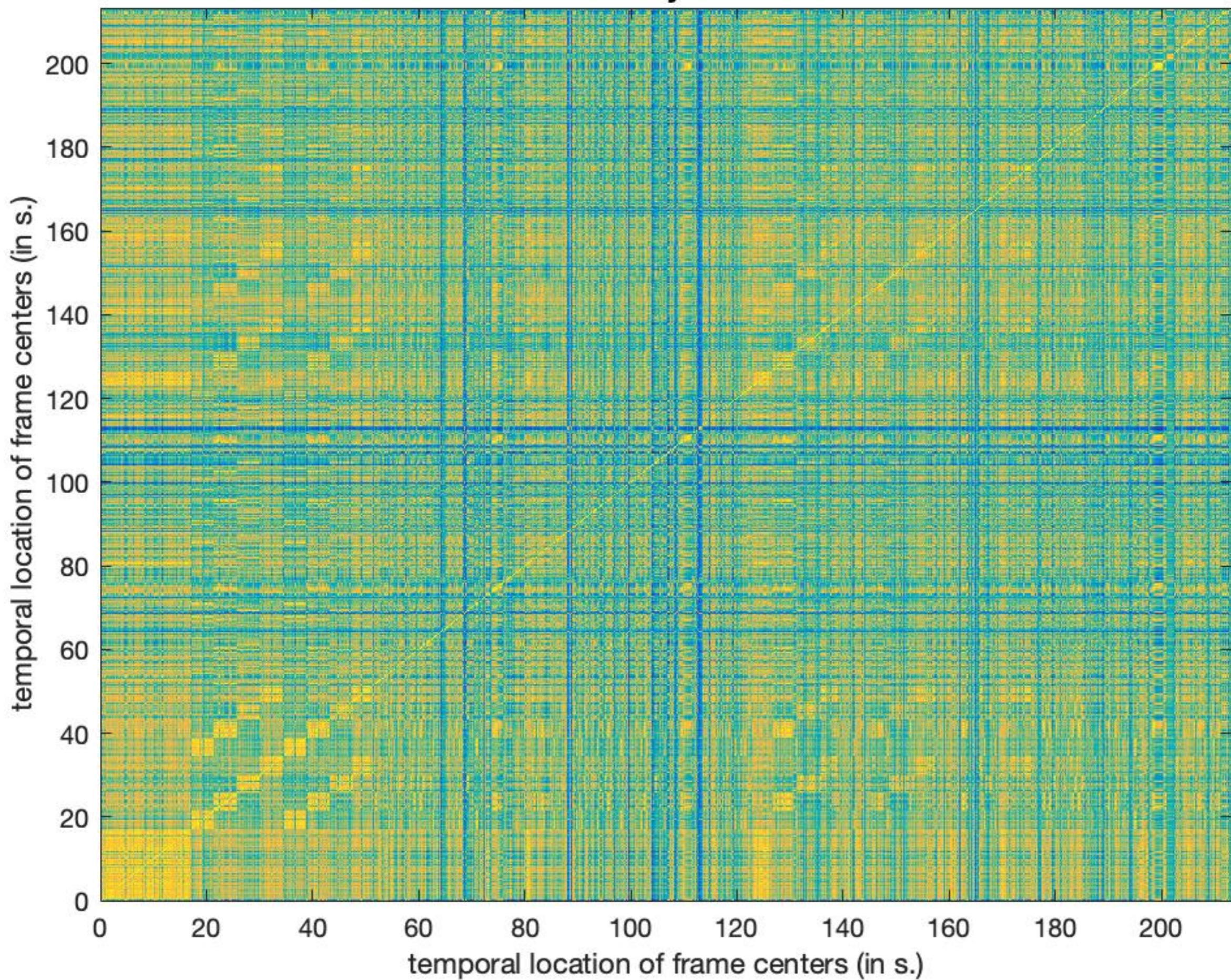
Musical Feature(s)
to
Musical Structure



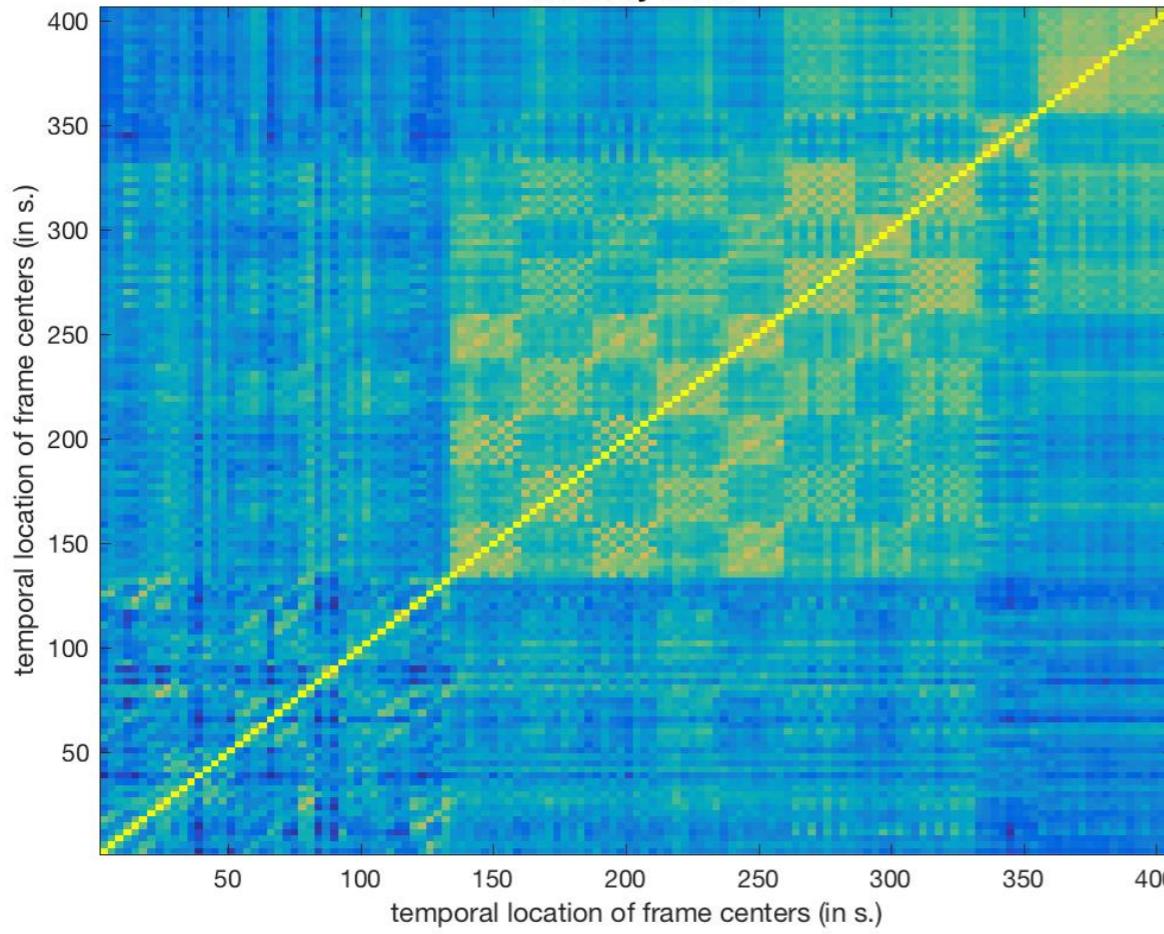
How does the structure of this part of the piece look like?



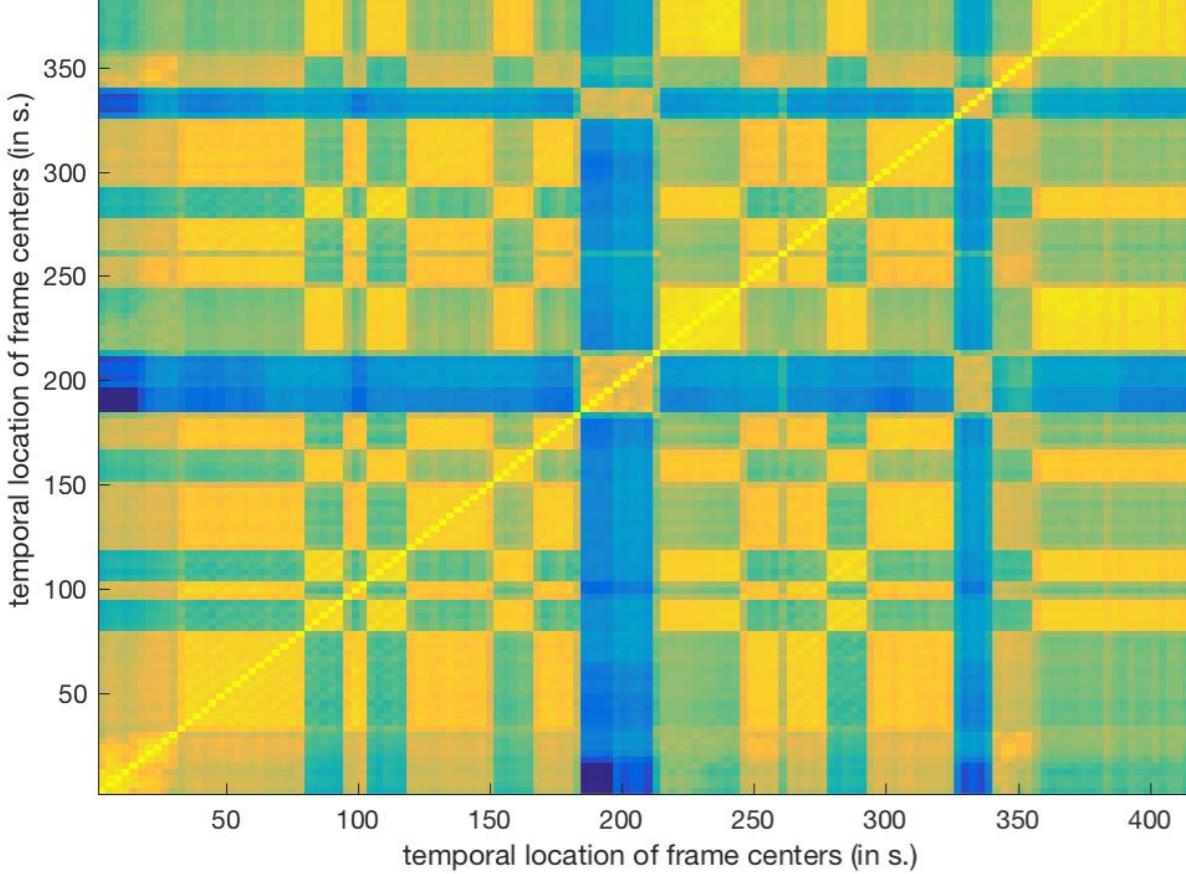
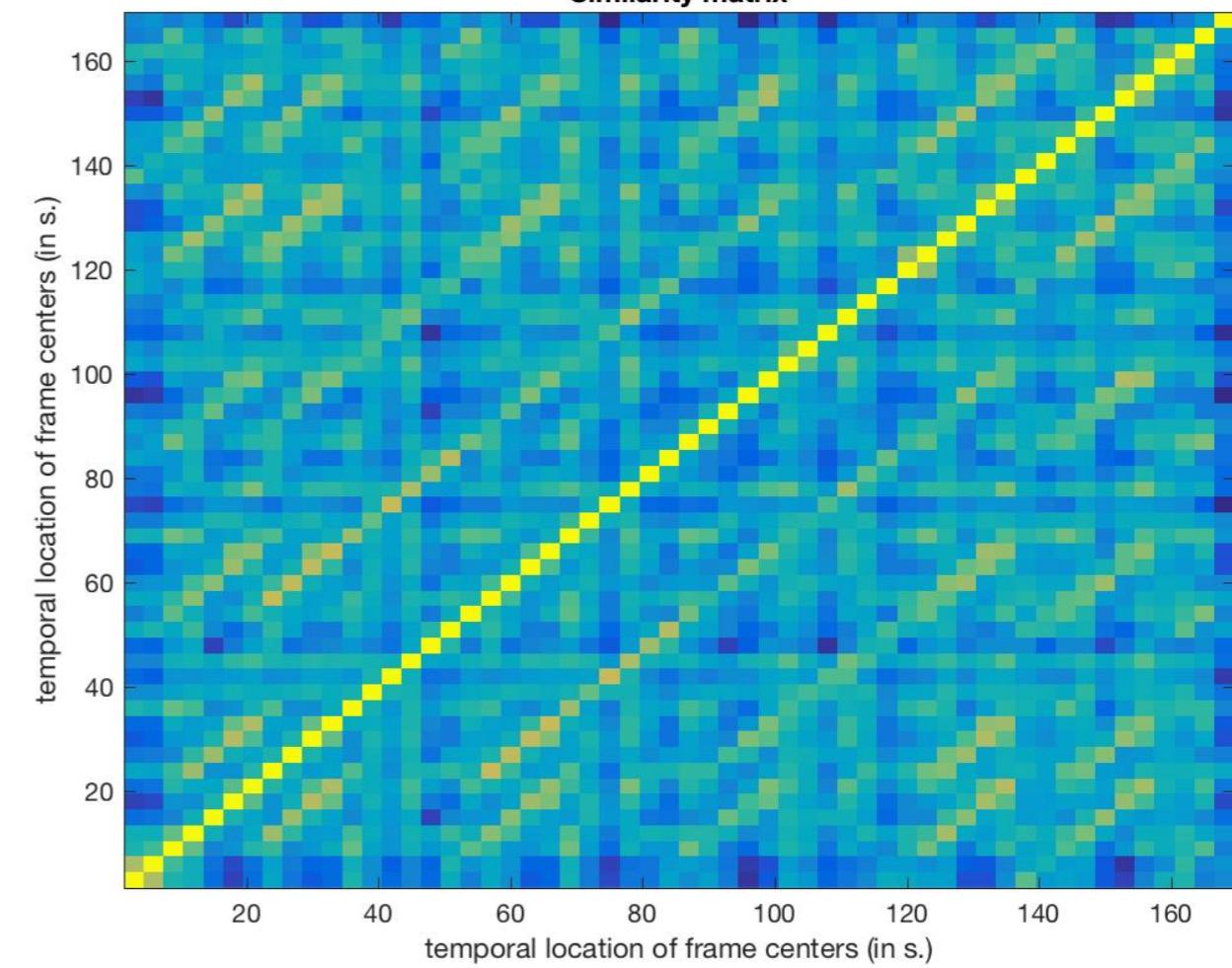
Similarity matrix

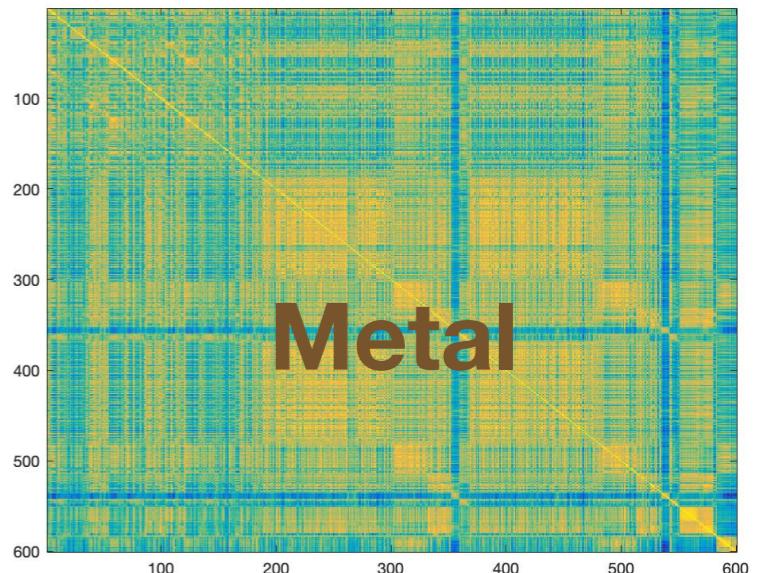
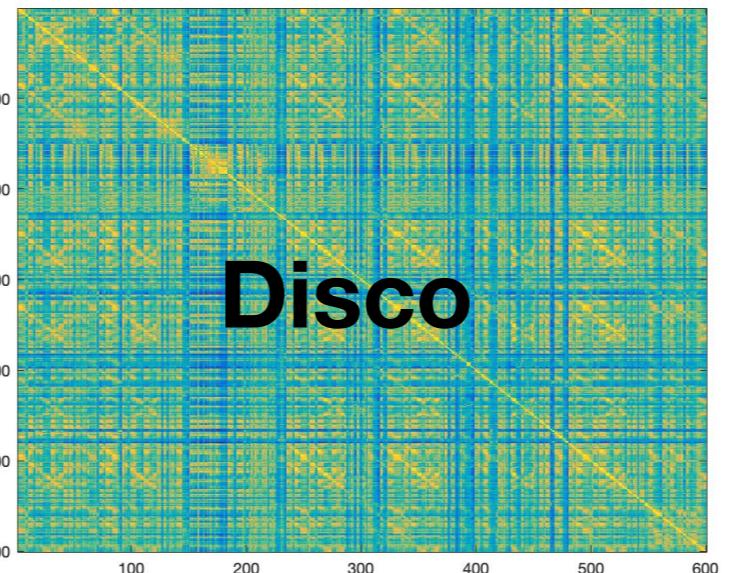
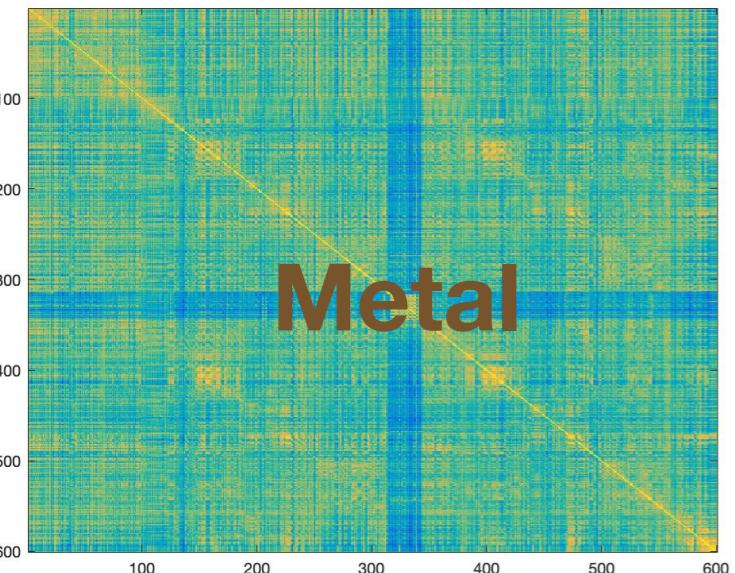
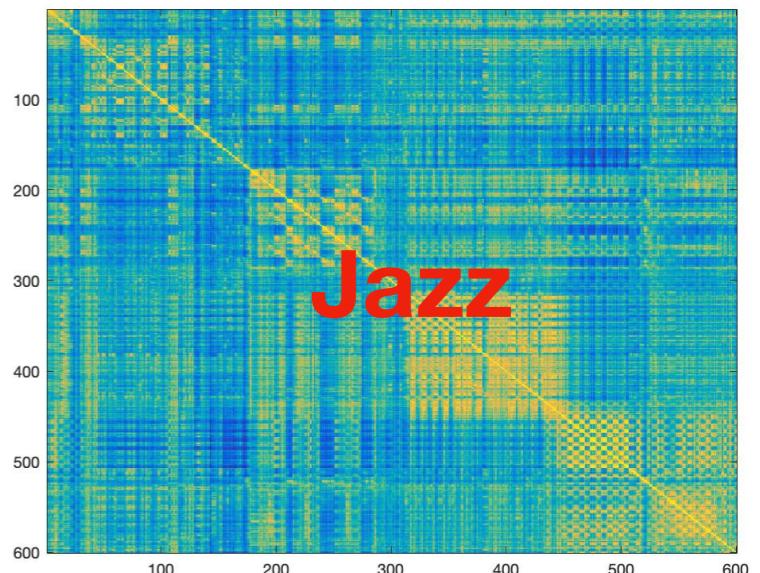
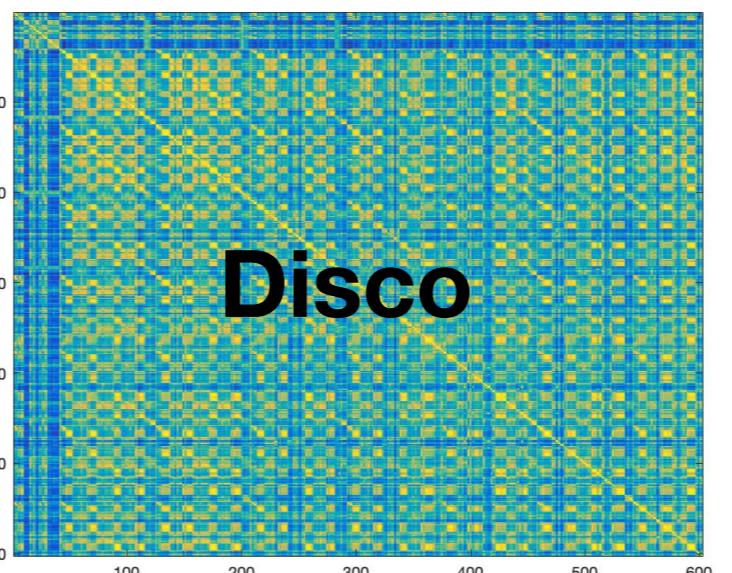
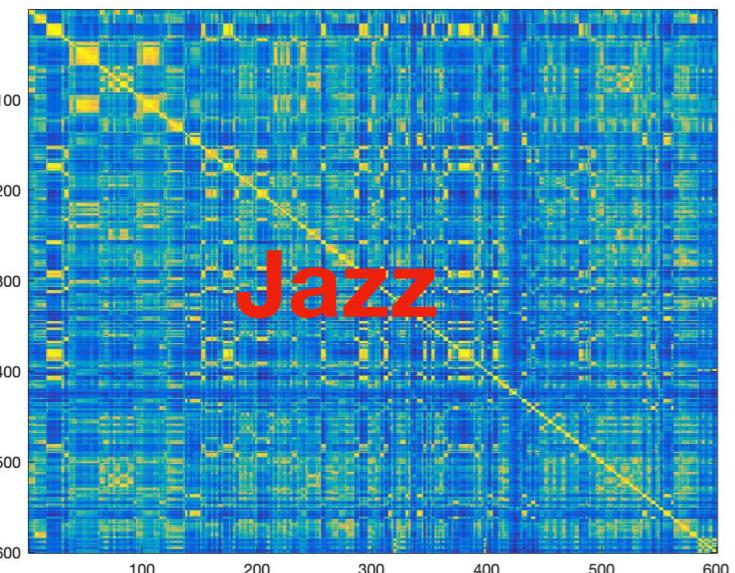
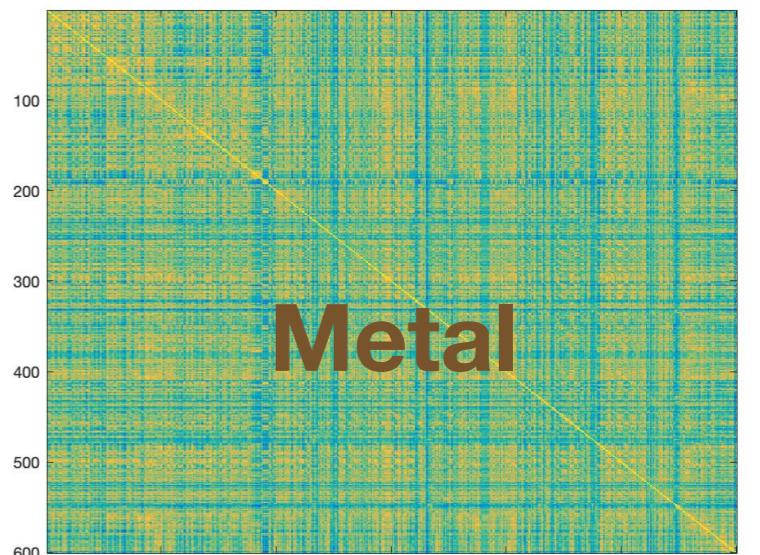
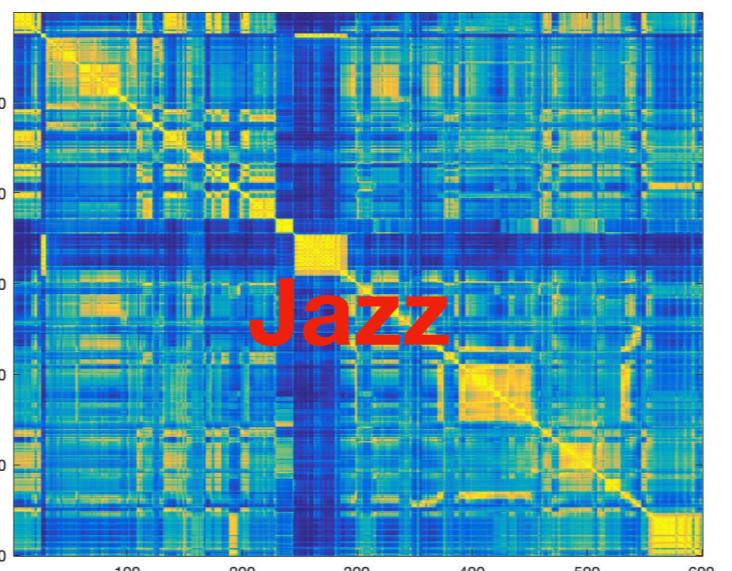
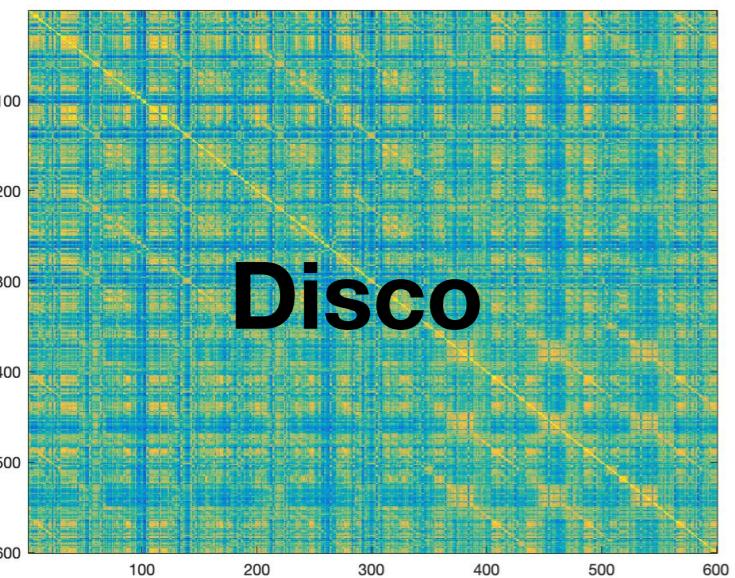


Similarity matrix



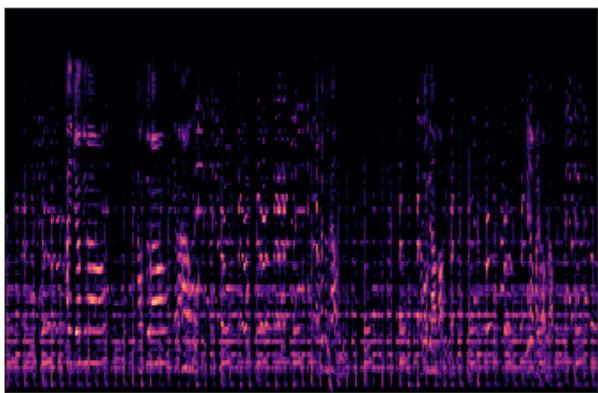
Similarity matrix



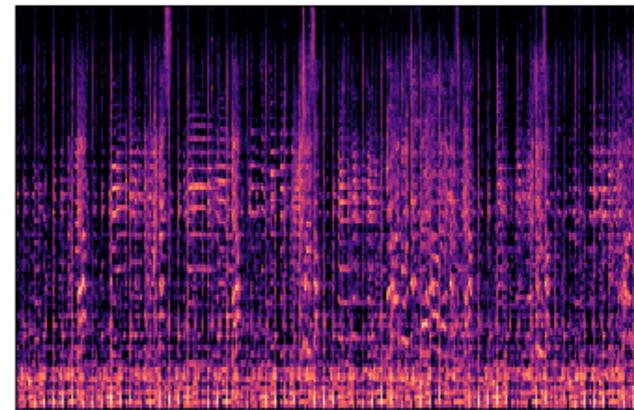


Genre Classification

can lyrics-based similarity matrices be used to identify genres ?



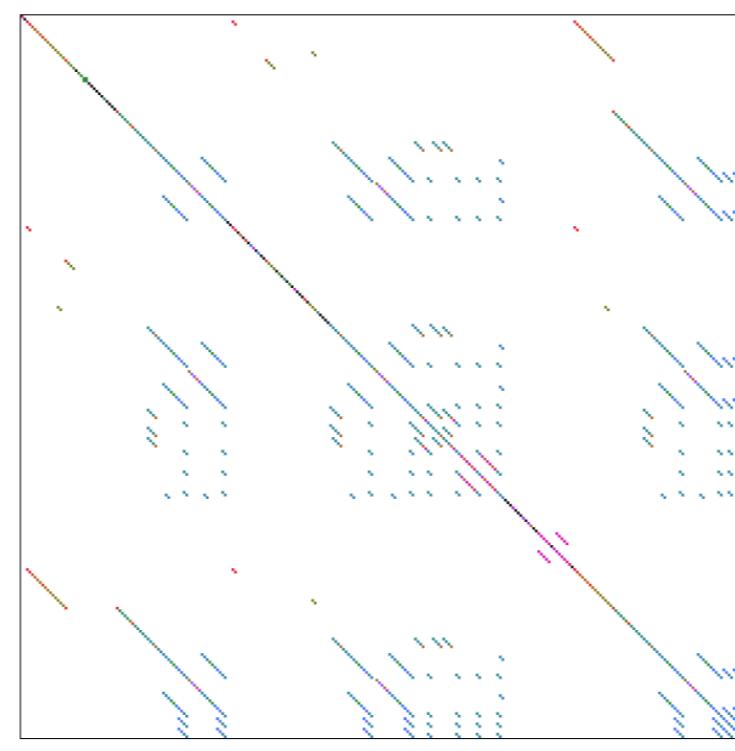
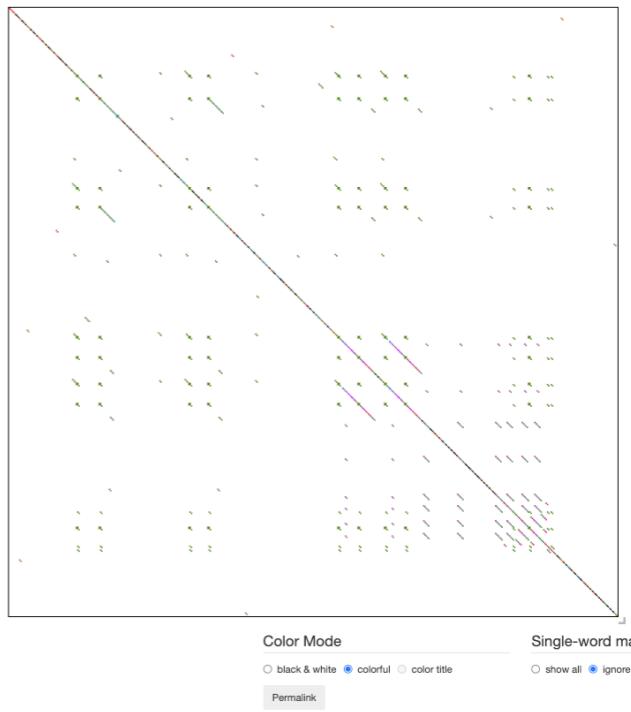
Genre 1



Genre 2

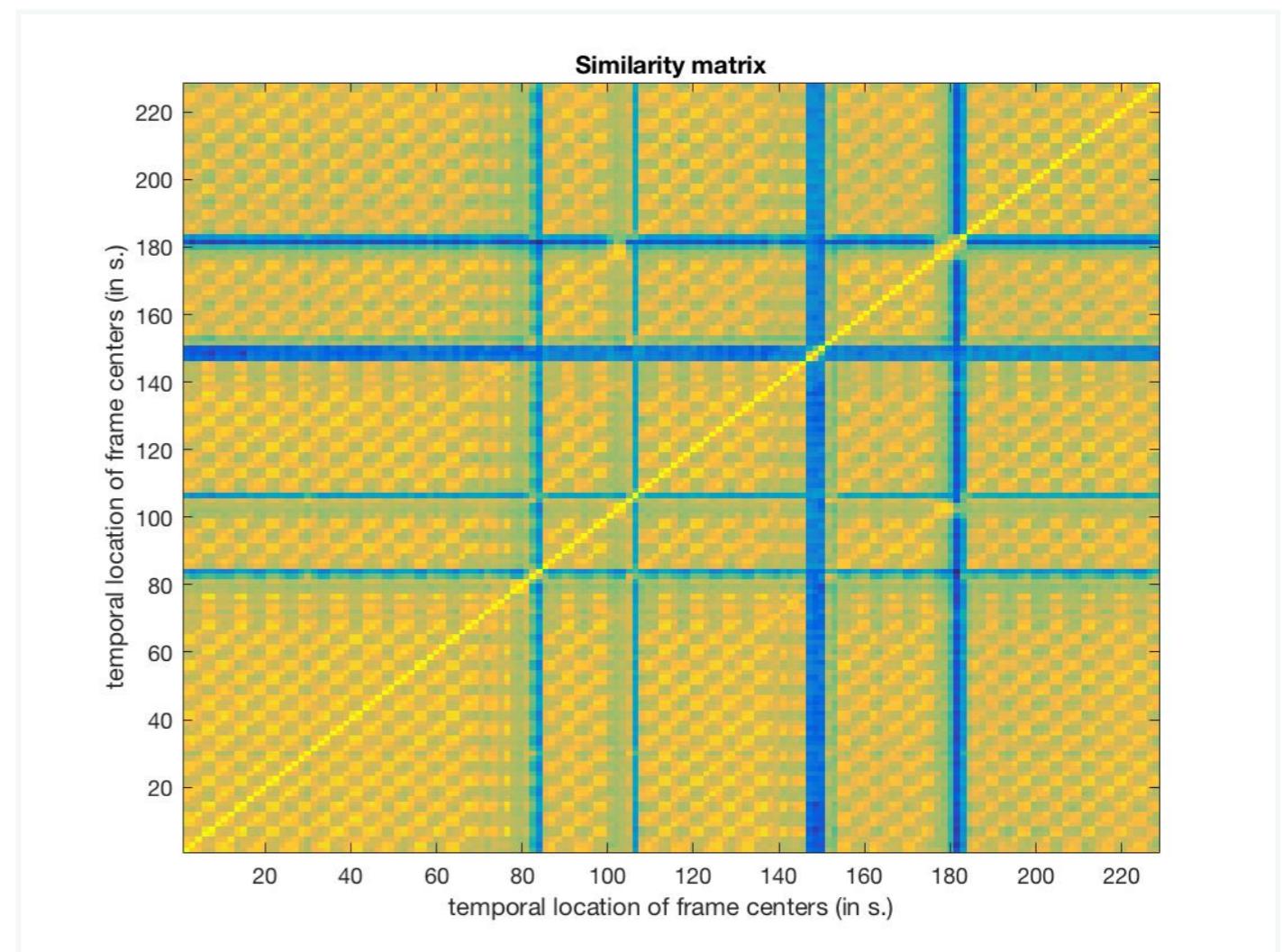
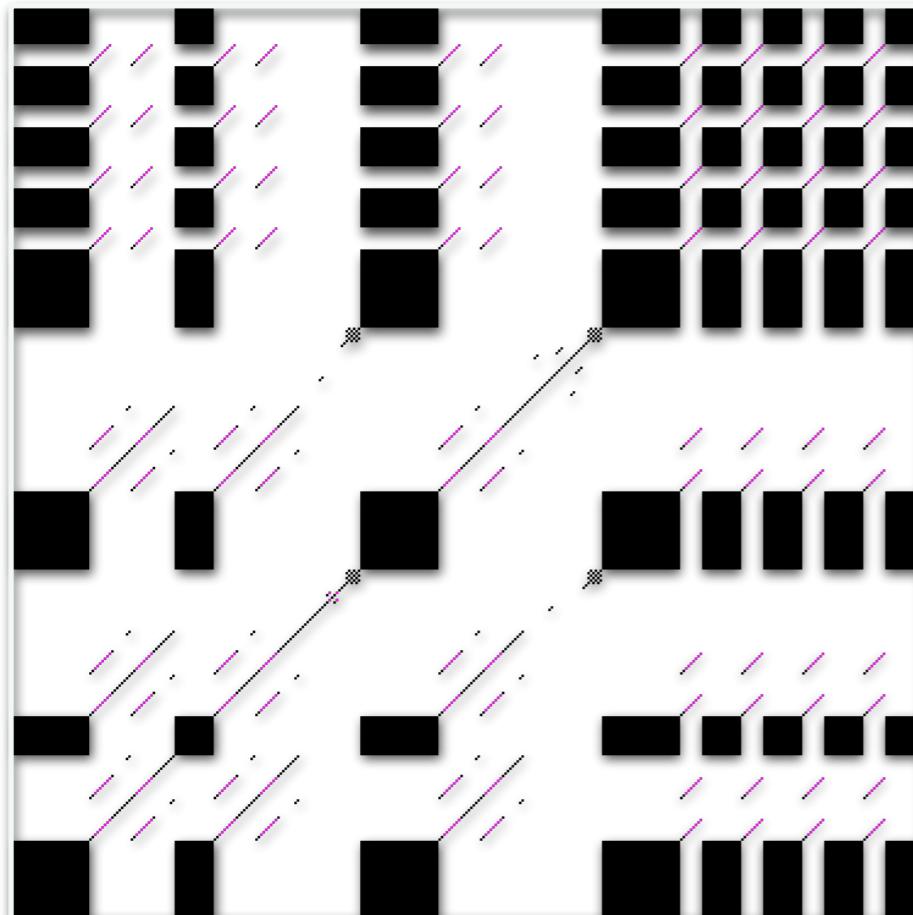
Visualizing Lyrical Structure

Vox



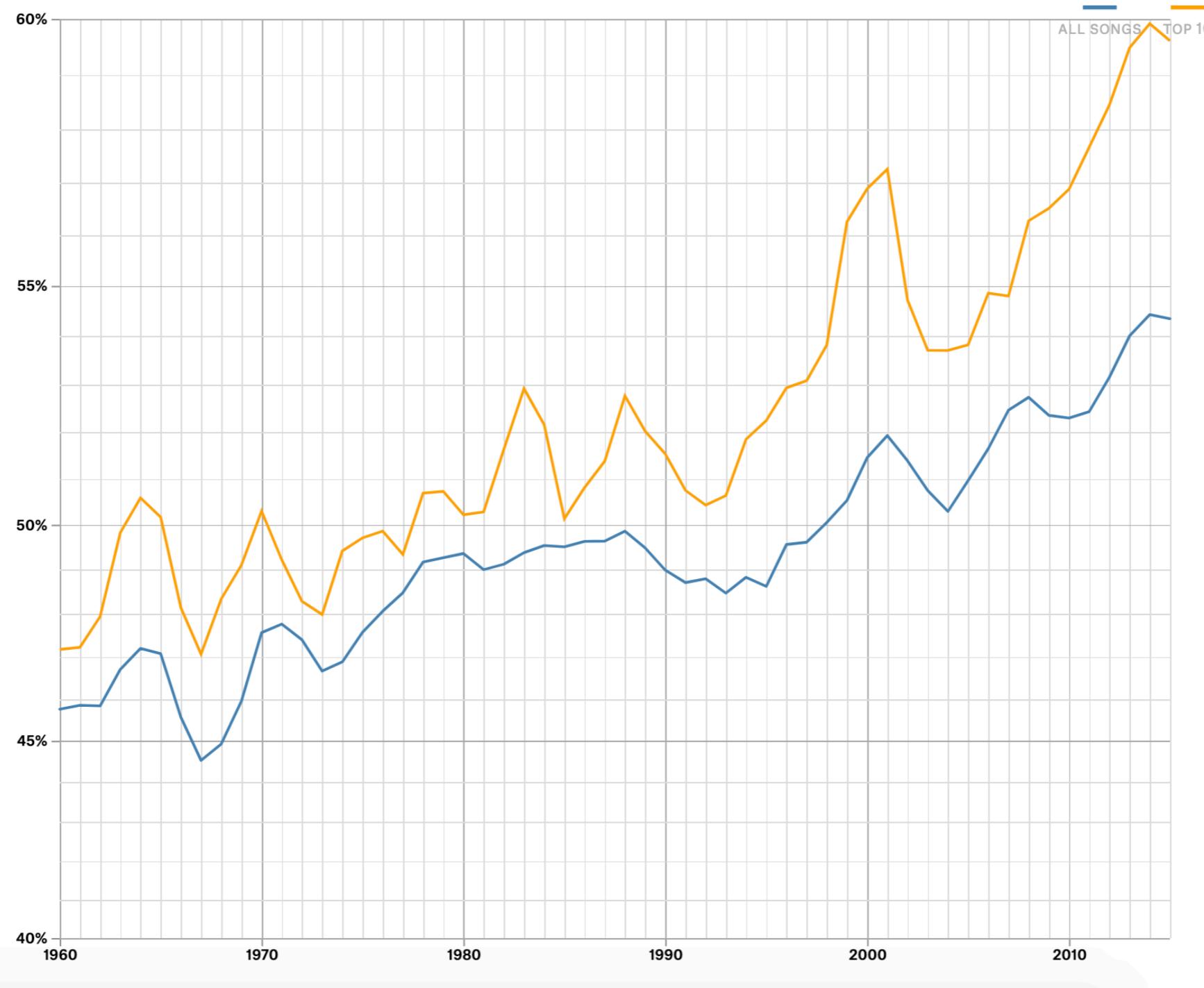
Custom			
Agg Lavan Majboori Nu Aan Jaan Di Pasoori Nu Zehar Bane Haan Teri Pee Jaavan Main Poori Nu Aana Si Oh Nahi Aaya Dil Bang Bang Mera Takraya Kaga Bol Ke Dass Jaave Paawan Gheyoe Dee Choori Nu Ranwan Ch Banwan Ch Oh Nu Lukavan Koyi Mainu Na Roke Tainu Khabar Kivein Hove Aa Jaave Dil Tera Poora Vi Na Hove Haan Baniya Banaiyan Di Galbaat Kivein Hove Aa Jaave Dil Tera	Poora Vi Na Hove Bhool Gayi Majboori Nu Dunnya Di Dastoori Nu Sath Tera Hai Bathera Poora Kar Zaroori Nu Aana Si Oh Nai Aaya Raasta Na Dikhiyaa Dil Humara De Sahara Khwahishat Adhoori Nu Waari Main Jaavan Main Tainu Bulawan Gall Saari Taan Hove Mere Dhol Judaiyan Di Tainu Khabar Kivein Hove Aa Jaave Dil Tera Poora Vi Na Hove Haan Baniya Banaiyan Di Galbaat Kivein Hove Aa Jaave Dil Tera	Aa Jaave Dil Tera Poora Vi Na Hove Mere Dhol Judaiyan Di Sardaari Na Hove Mere Dhol Judaiyan Di Aana Si Oh Nai Aaya Raasta Na Dikhiyaa Dil Humara De Sahara Khwahishat Adhoori Nu Waari Main Jaavan Main Tainu Bulawan Gall Saari Taan Hove Mere Dhol Judaiyan Di Tainu Khabar Kivein Hove Aa Jaave Dil Tera Poora Vi Na Hove Haan Baniya Banaiyan Di Galbaat Kivein Hove Aa Jaave Dil Tera	Agg Lavan Majboori Nu Aan Jaan Di Pasoori Nu Zehar Bane Haan Teri Pee Jaavan Main Poori Nu Ranwan Ch Banwan Ch Oh Nu Lukavan Koyi Mainu Na Roke Mere Dhol Judaiyan Di Tainu Khabar Kivein Hove Aa Jaave Dil Tera Poora Vi Na Hove Haan Baniya Banaiyan Di Galbaat Kivein Hove Aa Jaave Dil Tera

Repetition in Music

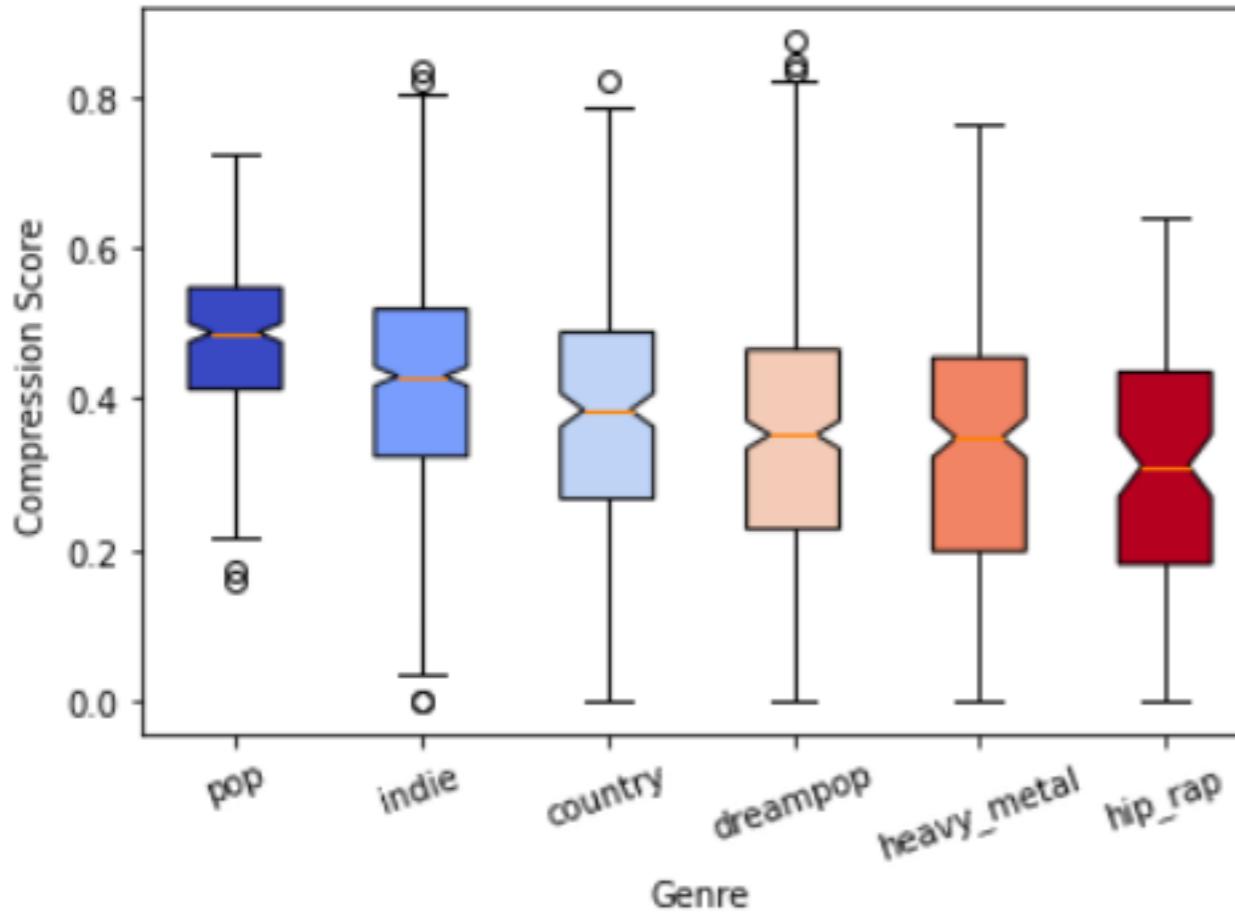


Repetition in Music

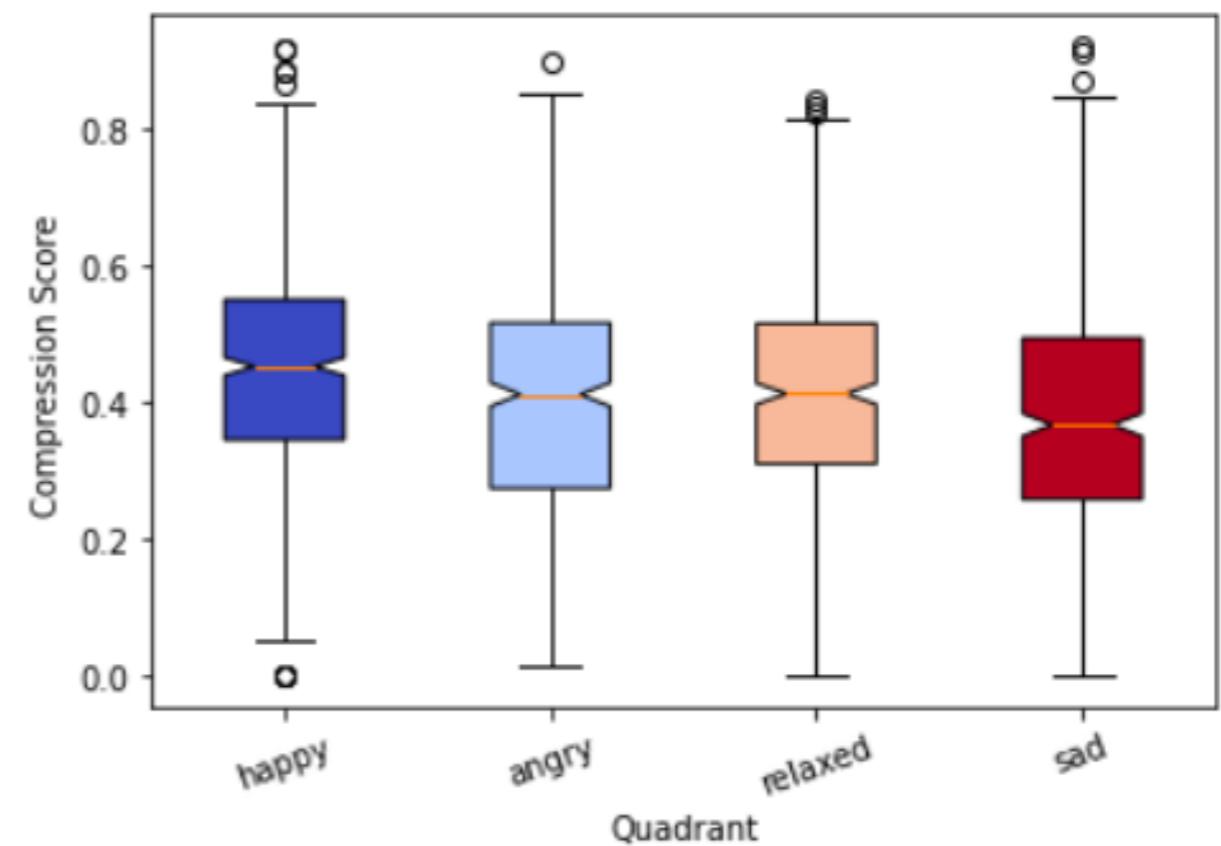
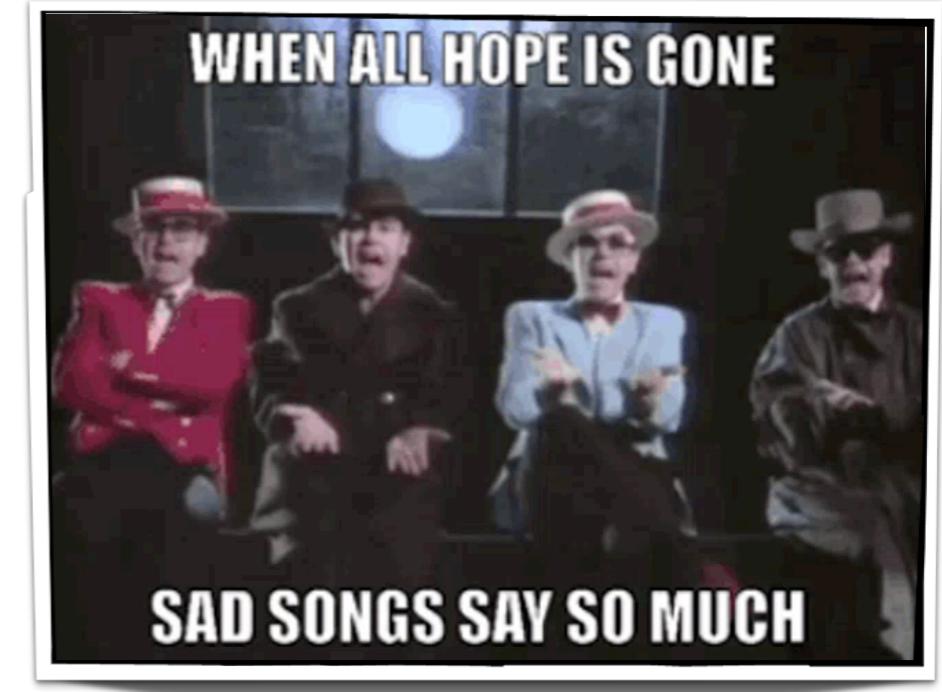
Repetition of Popular Music, by Year



Compressibility in Music



Which genres can be separated based on lyrical structure?



Genre Classification Demo

IEEE TRANSACTIONS ON SPEECH AND AUDIO PROCESSING, VOL. 10, NO. 5, JULY 2002

293

Musical Genre Classification of Audio Signals

George Tzanetakis, *Student Member, IEEE*, and Perry Cook, *Member, IEEE*

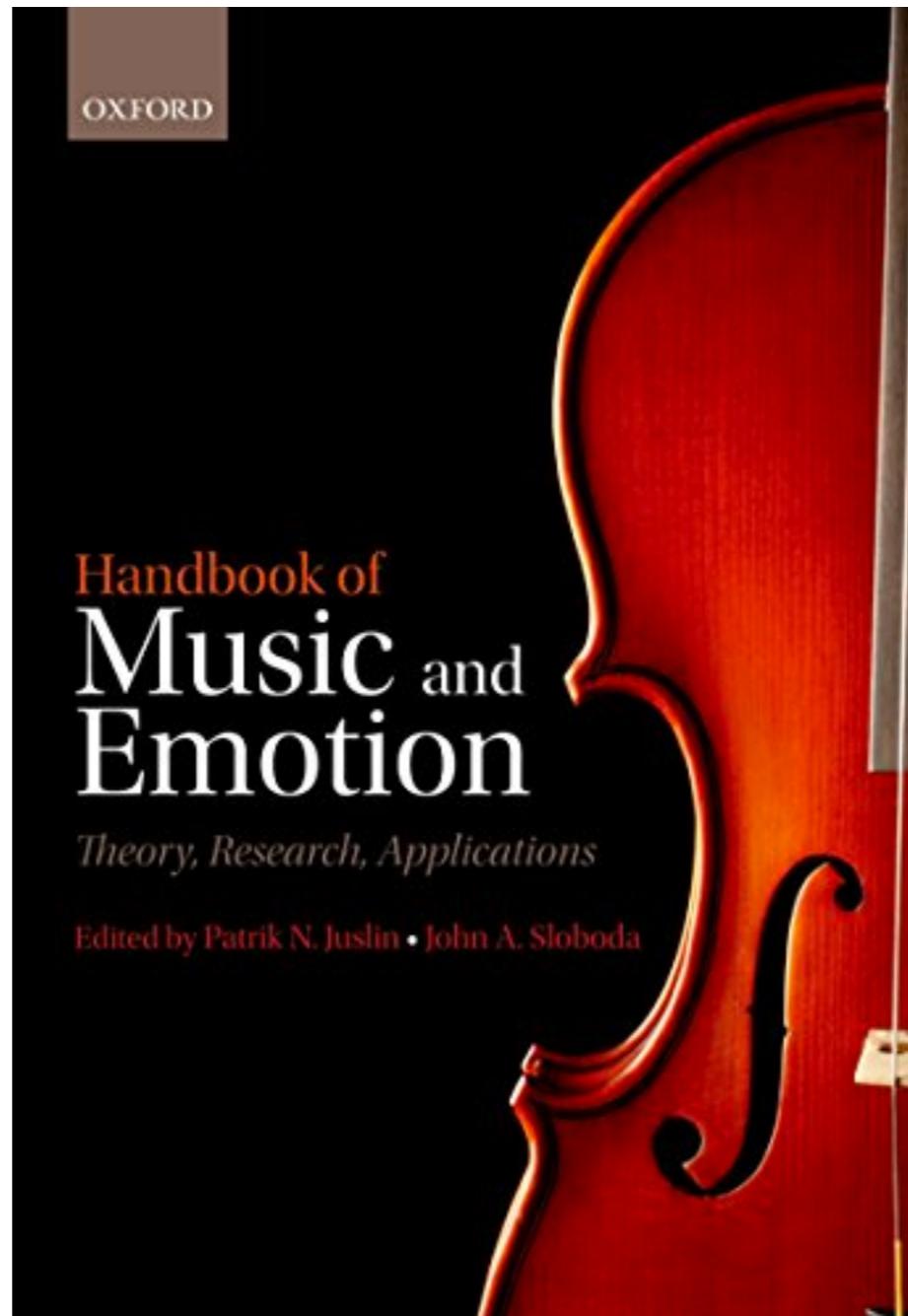
Emotion Classification

OXFORD

Handbook of
Music and
Emotion

Theory, Research, Applications

Edited by Patrik N. Juslin • John A. Sloboda





Emotions? Emotional Space?

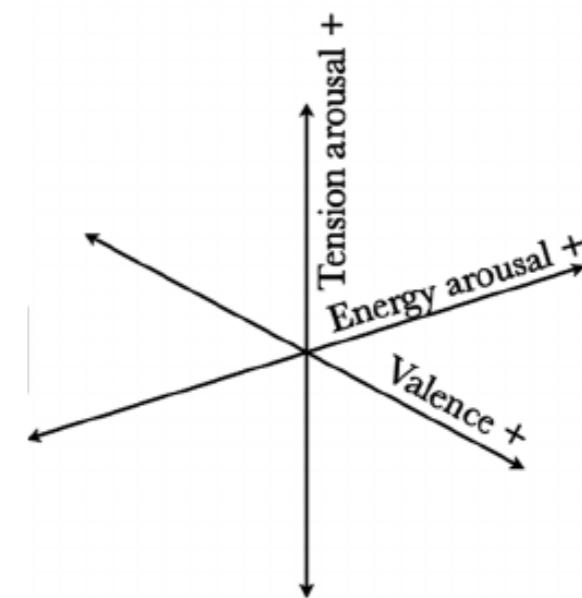
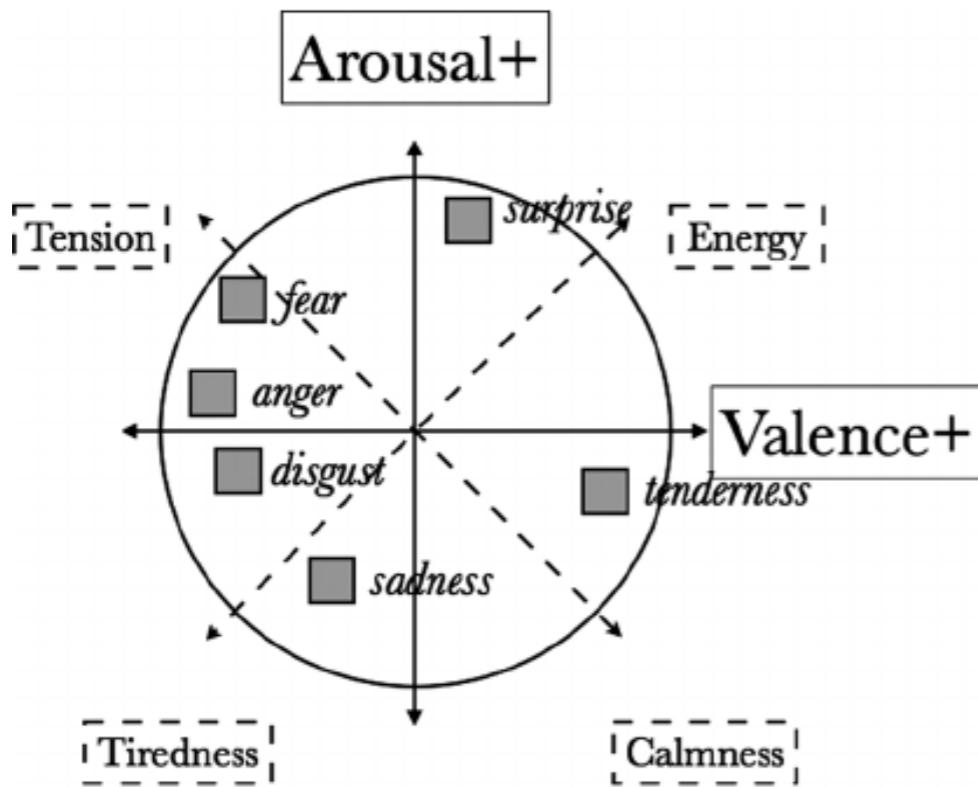


What emotions can music express?

Table 3. Frequencies with which various emotion labels were selected in response to the question “What emotions can music express?” ($N = 141$).

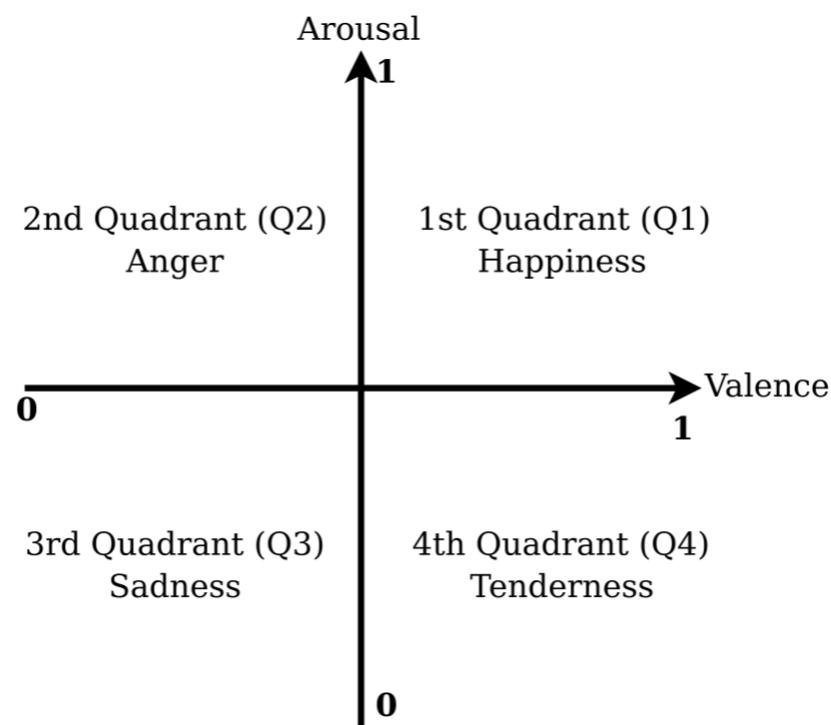
Emotion	Freq.	Emotion	Freq.	Emotion	Freq.
Joy	99% (98%)	Pride	71% (69%)	Curiosity	46% (63%)
Sadness	91% (91%)	Pain	70% (86%)	Boredom	45% (47%)
Love	90% (89%)	Desire	69% (74%)	Disappointment	43% (49%)
Calm	87% (89%)	Hope	67% (70%)	Guilt	42% (43%)
Anger	82% (83%)	Nostalgia	67% (76%)	Satisfaction	42% (57%)
Tenderness	82% (86%)	Fear	63% (79%)	Admiration	37% (37%)
Longing	77% (71%)	Contempt	55% (53%)	Jealousy	35% (42%)
Solemnity	76% (73%)	Tiredness	55% (52%)	Sympathy	34% (39%)
Anxiety	75% (90%)	Regret	53% (56%)	Shame	31% (39%)
Hate	74% (69%)	Expectancy	51% (66%)	Trust	30% (33%)
Humour	74% (87%)	Confusion	49% (65%)	Interest	29% (44%)
Loneliness	73% (79%)	Disgust	47% (51%)	Humiliation	28% (31%)
Tension	72% (89%)	Surprise	47% (68%)	Other	10% (16%)

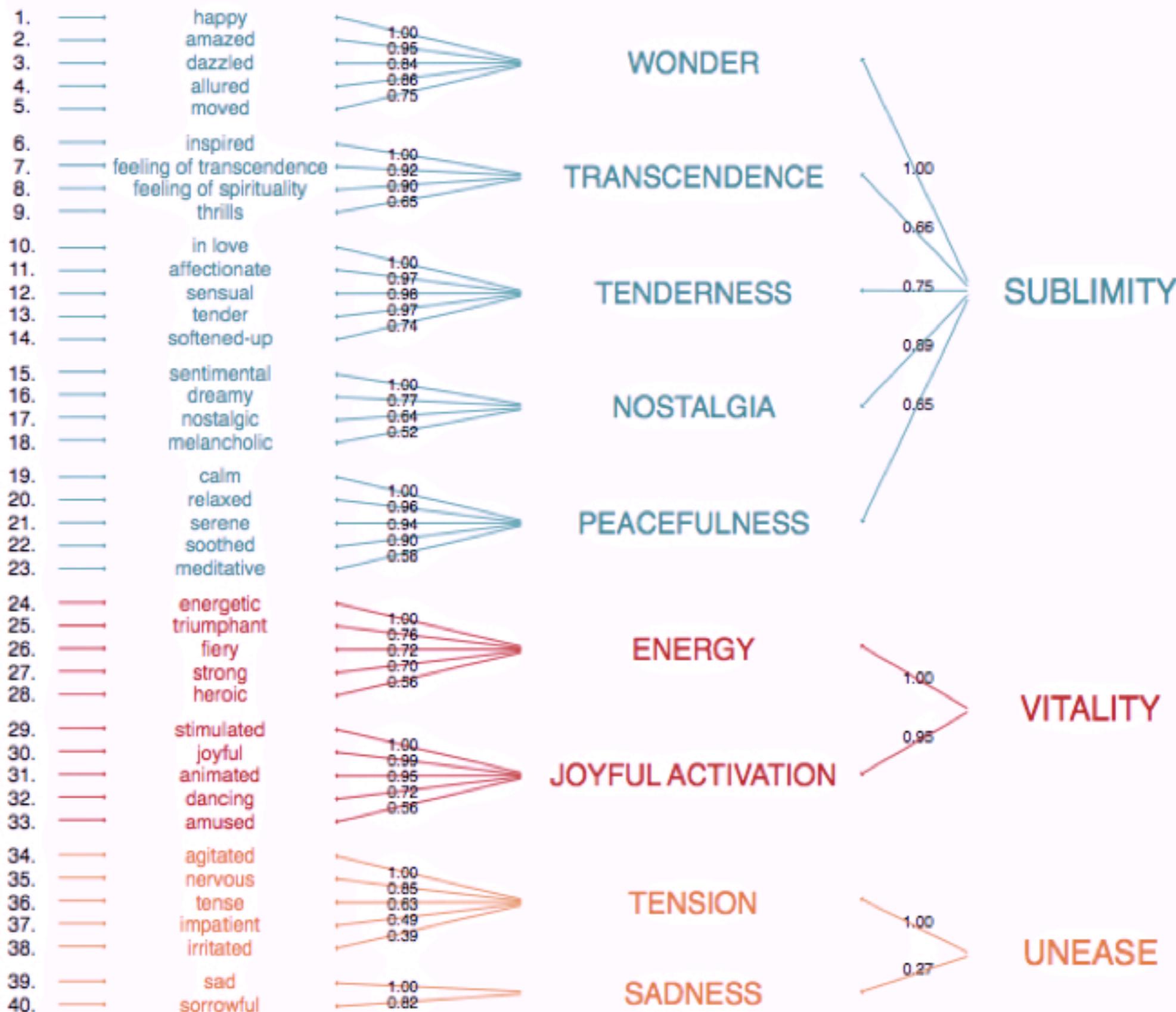
(Juslin & Laukka, 2004)



Schimmack & Grob model

Russell's circumplex model —
Thayer's model - - -
Basic emotion terms ■





The Geneva Emotional Music Scales (GEMS)

Theory

Emotions

A Discrete

Happy, sad, anger, fear, disgust (e.g., Ekman, 1992)

B Dimensional

Activity & valence (Russell, 1980)

Tension & energy (Thayer, 1989)

C Aesthetic

joy	happiness	joy
peaceful	love	beautiful
spirituality	surprise	amusing
tenderness	calm	energetic
wonder	awe	dreamy
nostalgia	interest	triumphant
power	nostalgia	calm
tension	pride	erotic
sad	disgust	anxious
Zentner et al (2008)	Juslin et al (2011)	Cowen et al (2020)



How is music able to
communicate emotional
meaning?



How can music **induce**
emotions?

How is music able to
communicate emotional
meaning?

.... in the absence of lyrics?

Cues used to express (and infer) emotions in other human domains

- **Speech and vocal expression of emotion** (Juslin & Laukka, 2003)
 - Music communicates emotional meaning to listeners by exploiting the acoustic code for vocal expression of emotions
 - Many similarities between musical and vocal expression of emotion in terms of acoustic features, recognition accuracy

Cues used to express (and infer) emotions in other human domains



potential problems: cultural differences, mild/subtle cues, contextual & personal factors

Musical features



- Distinct structural and performance features contribute to the expression of different emotions
 - **Sadness:** ?
 - **Happiness:** ?
 - **Anger:** ?
 - **Tenderness:** ?

Musical features



- Distinct structural and performance features contribute to the expression of different emotions
 - **Sadness:** minor mode, low tempo, low pitch, narrow melodic range, low sound level, little sound level variability, soft timbre, and legato articulation
 - **Happiness:** major mode, regular rhythm, fast tempo, high pitch, wide melodic range, consonant harmonies, and staccato articulation
 - **Anger:** fast tempo, high sound level & sound level variability, low key clarity/atonality, fast tone attacks, and microstructural irregularity
 - **Tenderness:** major mode, slow tempo, low sound level, little sound level variability, low pitch level, little pitch variability, smooth rhythm, and slow tone attacks

Musical features



- Distinct structural and performance features contribute to the expression of different emotions
 - **Sadness:** minor mode, low **tempo**, low pitch, narrow melodic range, low **sound level**, little sound level variability, soft timbre, and legato articulation
 - **Happiness:** major mode, regular rhythm, fast **tempo**, high pitch, wide melodic range, consonant harmonies, and staccato articulation
 - **Anger:** fast **tempo**, high **sound level** & sound level variability, low key clarity/atonality, fast tone attacks, and microstructural irregularity
 - **Tenderness:** major mode, slow **tempo**, low **sound level**, little sound level variability, low pitch level, little pitch variability, smooth rhythm, and slow tone attacks

The image shows a dual-screen interface. On the left screen, a weather application displays a yellow smiley face icon in the center, surrounded by a gradient background of orange, yellow, green, and red. It also features small icons for lightning, sun, and moon in the corners. On the right screen, a music player application is open. At the top, there are standard media controls: back, forward, play/pause, volume, and track information showing "I'll Bru 0:00". Below these are sections for "KIRJASTO" and "STORE". The "KIRJASTO" section lists categories like Musiikki, Elokuvat, TV-ohjelmat, Podcastit, iTunes U, Kirjat, Ohjelmat, and Radio, with counts of 9 and 11 respectively. The "STORE" section shows a track titled "Toistettava" by Bruce Springsteen, featuring a photo of him and the text "RG", "BRUCE SPRINGSTEEN", and "MAGIC". The bottom of the screen shows additional music player controls (+, -) and a track count of "8 kappalet".

I'll
Bru
0:00

KIRJASTO

- Musiikki
- Elokuvat
- TV-ohjelmat
- Podcastit
- iTunes U
- Kirjat
- Ohjelmat
- Radio

9 11

STORE

Toistettava

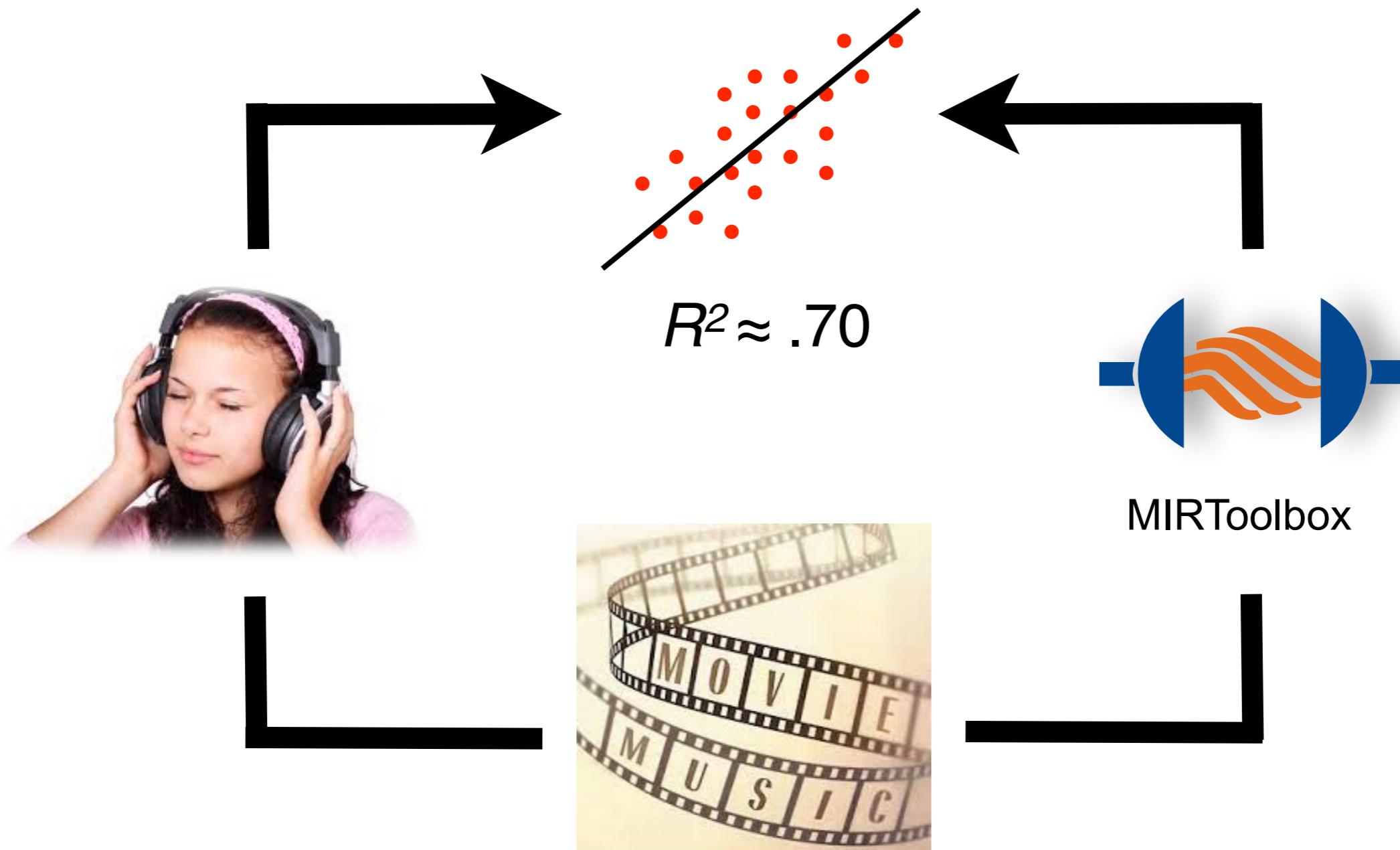
RG

BRUCE SPRINGSTEEN

MAGIC

+ - ↻ ↻ 8 kappalet

MIR & emotion prediction



Eerola, T., Lartillot, O., & Toiviainen, P. (2009) **Prediction Of Multidimensional Emotional Ratings In Music From Audio Using Multivariate Regression Models**. In proc. of 10th International Society for Music Information Retrieval Conference (ISMIR).



basic emotion
concepts and
dimensional ratings,
on Likert scales

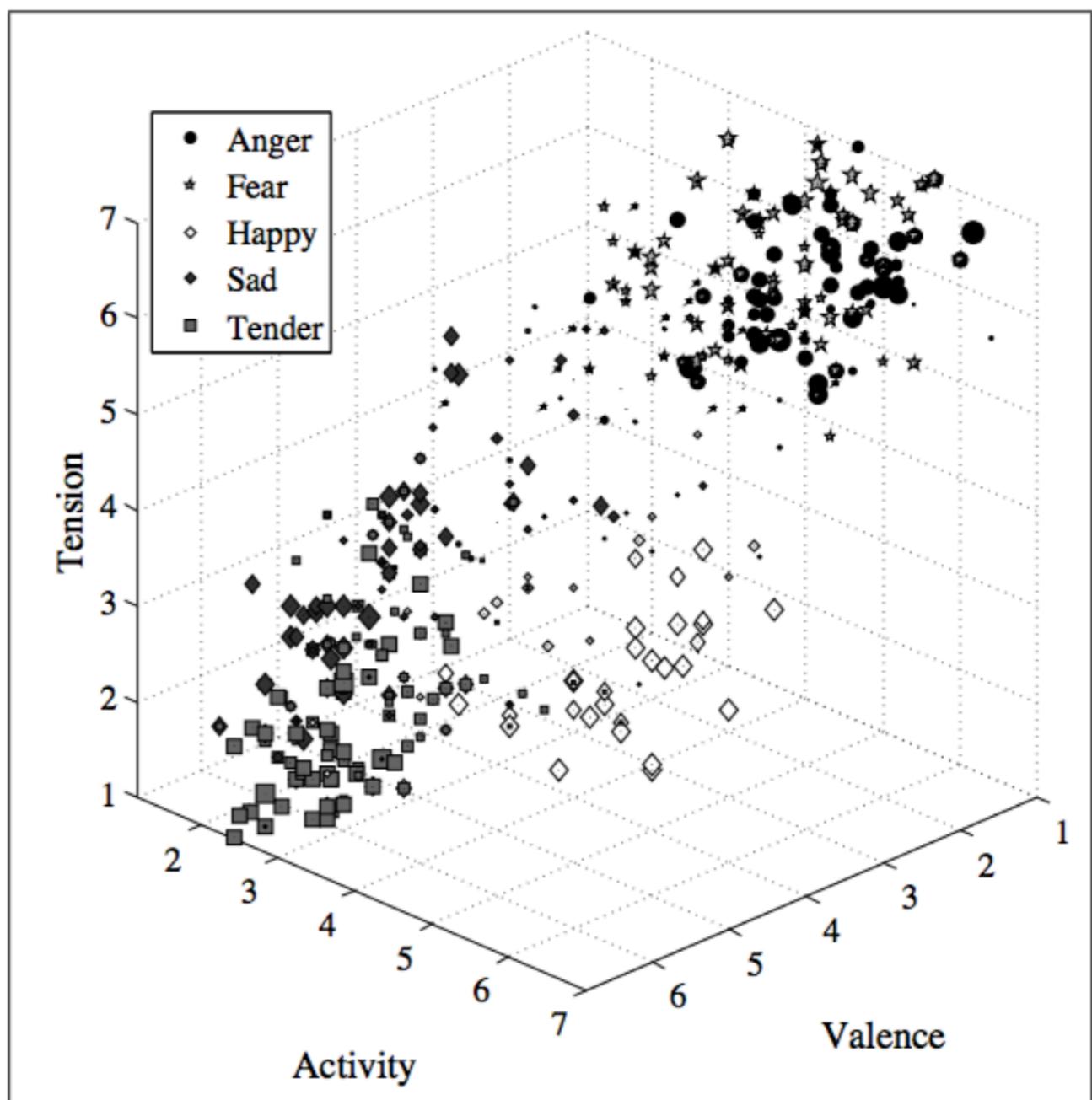
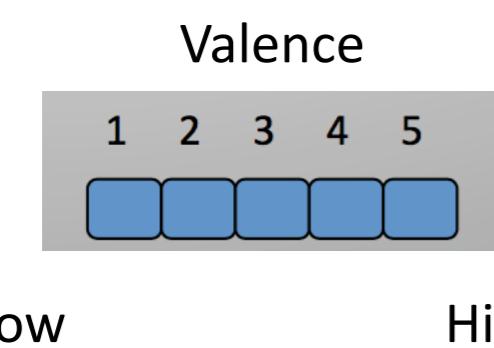
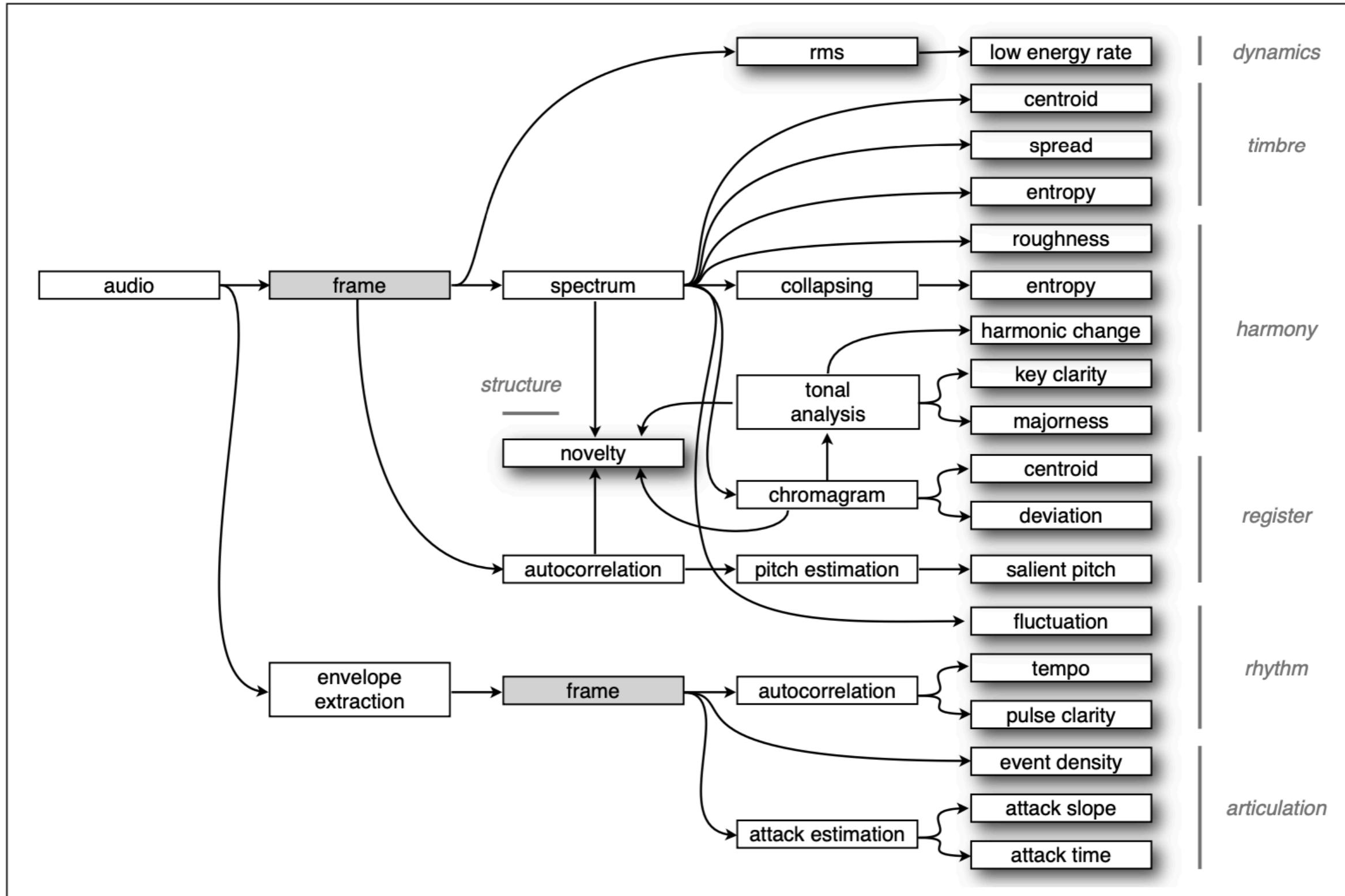


Figure 1. Average ratings of the three dimensions and basic emotions for the 360 soundtrack excerpts.



Eerola, T., Lartillot, O., & Toiviainen, P. (2009) **Prediction Of Multidimensional Emotional Ratings In Music From Audio Using Multivariate Regression Models.** In proc. of 10th International Society for Music Information Retrieval Conference (ISMIR).

	Prediction rate (R^2)		
Model	Valence	Activity	Tension
MLR	.64	.75	.67
PCA	.42	.74	.51
PLS	.70	.77	.71
MLR $_{\lambda}$.66	.74	.69
PCA $_{\lambda}$.51	.73	.63
PLS $_{\lambda}$.72	.85	.79

	Prediction rate (R^2)				
Model	Angry	Scary	Happy	Sad	Tender
MLR	.46	.55	.46	.38	.38
PCA	.66	.67	.60	.59	.54
PLS	.66	.62	.61	.61	.50
MLR $_{\lambda}$.56	.55	.63	.54	.45
PCA $_{\lambda}$.56	.47	.53	.52	.45
PLS $_{\lambda}$.70	.74	.68	.69	.58

Cues used to express (and infer) emotions
in other human domains

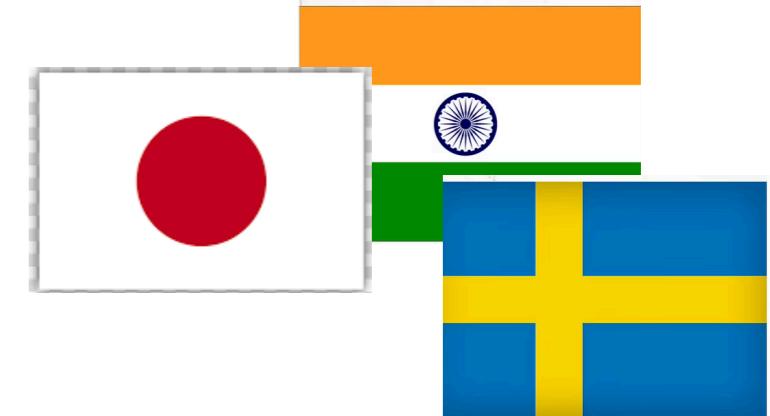


- **Human movement & gesture**
 - All sounds from traditional acoustics instruments are produced by human movement -> we can “hear” this movement in music
 - Music emulates the speed, posture, and smoothness/jerkiness of human movement and gestures when experiencing/expressing emotions (Jackendoff & Lerdahl, 2006)

Other cues

- **Culturally learned cues**
 - Mode (minor/major)
 - Pitch (high/low)
 - Contextual associations (e.g., wedding and funeral music)
 - Frequent pairing with narrative content (song lyrics, films, TV, opera)

Other cues



- **Culturally learned cues**

- better identification of basic emotions (anger, fear, happiness, and sadness) than non-basic ones (e.g., solemnity, humor, and longing)
- better able to identify the intended emotions in music from their own culture

peacefulness,
longing,
neutral solemnity,
affection,
happiness,
spirituality, fear,
humor, sadness,
anger,

Other cues

- **Human characteristics/virtual person**
 - Music is assigned attributions that normally would be assigned to a person -> music creates a “virtual person”? (Watt & Ash, 1998)

“.....a piece of music can be assigned an attribution that has the value female rather than male. This need not imply that the music is female, just that the attribution that is made has more of the quality of female rather than of the quality of male. These attributions are made to the music, not to the composer or the performer. Loosely speaking, music creates a virtual person.”

How is music able to **communicate** emotional meaning?



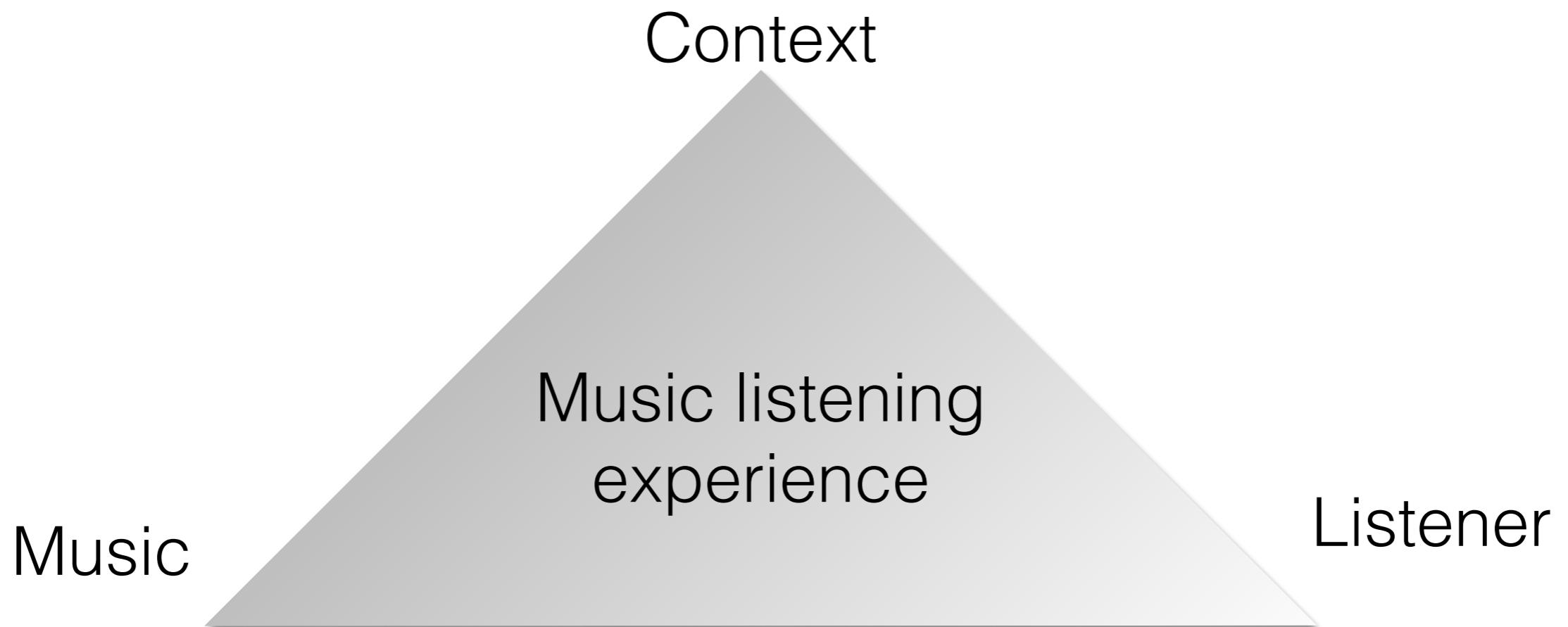
How can music
induce emotions?

Mechanisms:

- **B**rain stem reflexes
- **R**hythmic entrainment
- **E**valuative conditioning
- Emotional **C**ontagion
- **V**isual imagery
- **E**pisodic memory
- **M**usical expectancy
- **A**esthetic Judgement

Different mechanisms may be at function simultaneously, and lead to differing emotional responses (i.e., mixed emotions)

Relevance of context



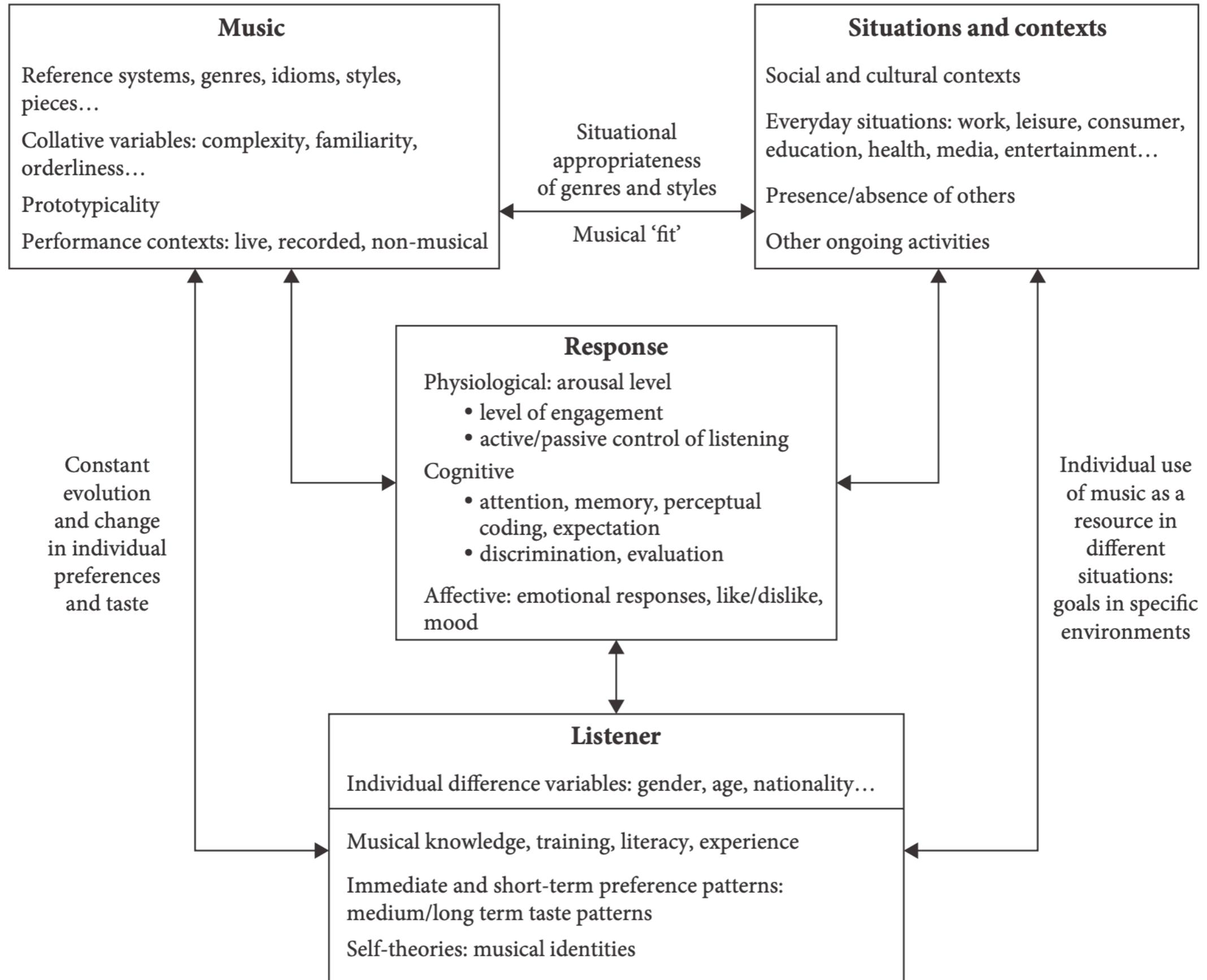


FIGURE 17.2 Reciprocal feedback model of musical response.



Classification Tasks

Genre & Emotion

<https://pudding.cool/2017/05/song-repetition/>