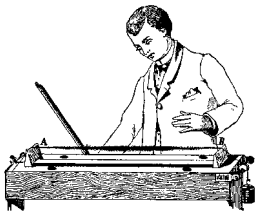


DIGITAL AUDIO EMOTIONS - AN OVERVIEW OF COMPUTER ANALYSIS AND SYNTHESIS OF EMOTIONAL EXPRESSION IN MUSIC

Anders Friberg

Speech, Music and Hearing, KTH, Stockholm,Sweden
http://www.speech.kth.se

Why music?

Our (unexpected?) emotional sensivity to sound:

•Audio is an important information carrier for speech , biological motion (humans, animals) as well as environmental motion (rocks, wheather…).

•Auditive and visual stimuli are ”overanalysed” in search for meaningful communication.

•Emotions can be triggered from violations of expectations in sound sequences (Huron 2006)

Music usage: get an emotional experience, mood controller, energy injection, motion inducer (dance)…

Definitions

Score features – features inherent in the score (notes, harmony, meter, mode…)

Performance features – features modified by the performer (tempo, dynamics, articulation…)

Emotion perception - listeners' perception of emotional expression

Emotion induction - listeners' own emotional reaction

Traditional research topics

- Which emotions can be reliably expressed in music?
- Which musical features influence the perceived emotional expression?

Recent research topics

- How can emotional expression be analysed and synthesized by the computer?
- Which applications can be constructed?

History of empirical research in music and emotion

Overview: Gabrielsson and Lindström 2001

Started in the 1900 century

~1930 Kate Hevner made a series of experiments with systematical manipulation of features

~ 1990 - Uppsala University with Gabrielsson and co-workers investigated performance features in detail

1997 Juslin's lens model for quantifying the relation between performer and listener using musical features.

~ 2000- Emotion models using modern data analysis tools

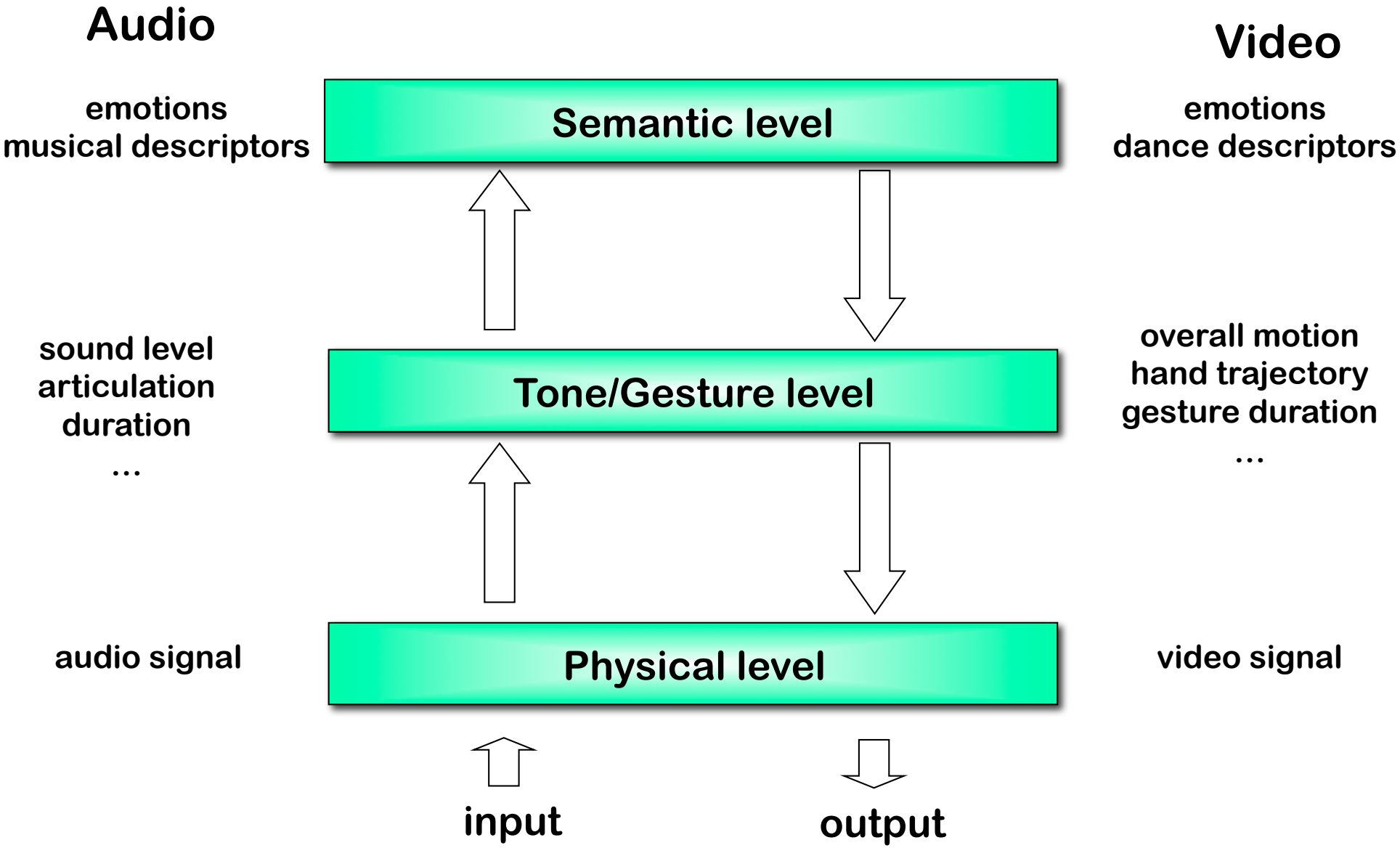
More recent

Music and Emotion, Juslin and Sloboda, eds., 2001

Gabrielsson and Lindström 2001 – historic overview with a summary of experimental methods and the impact of every musical feature (86 references)

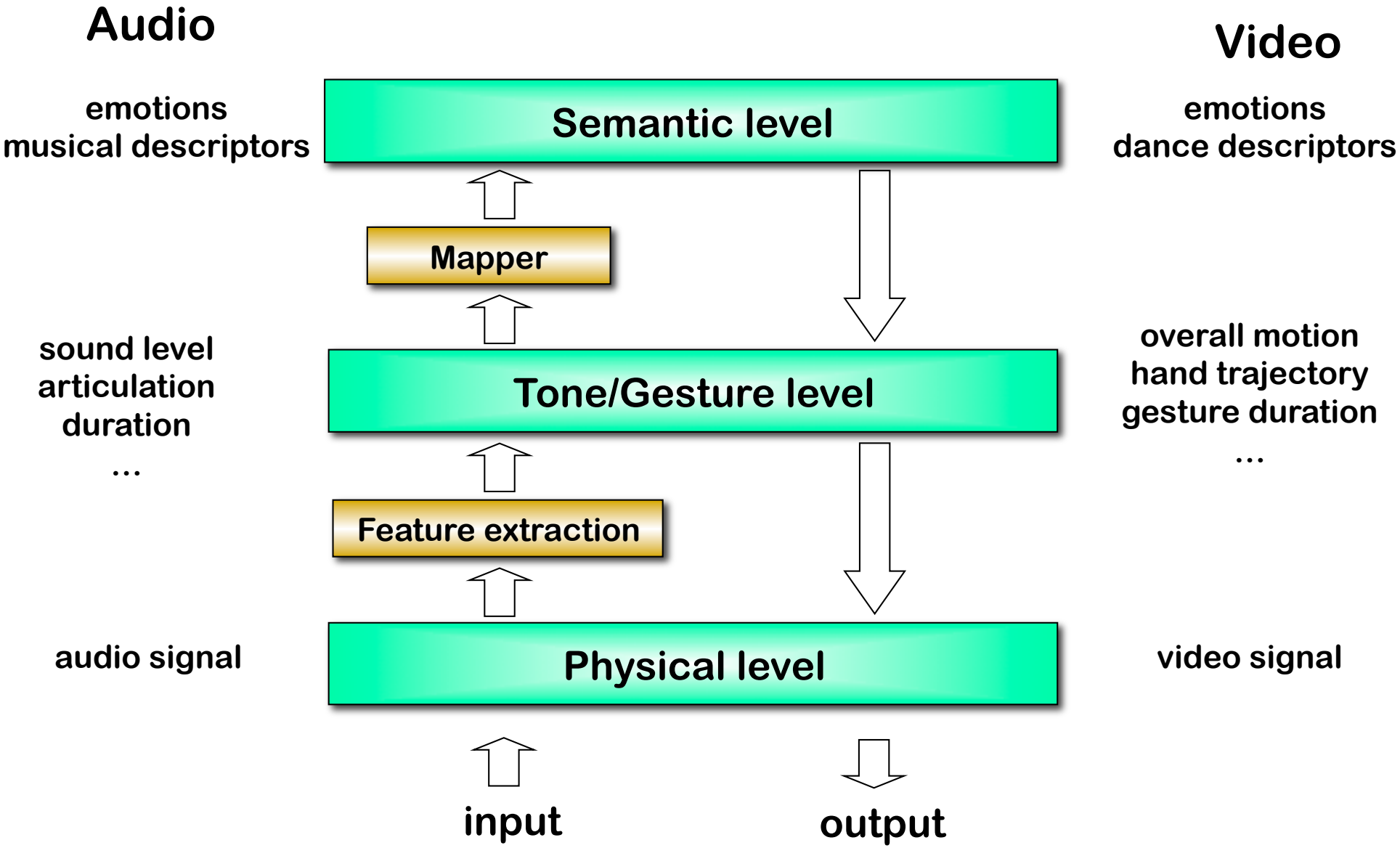
Juslin and Laukka 2003 - a meta-analysis of 41 papers of emotional expression in music performance and 104 articles studying emotions in speech

Framework for analysis and synthesis



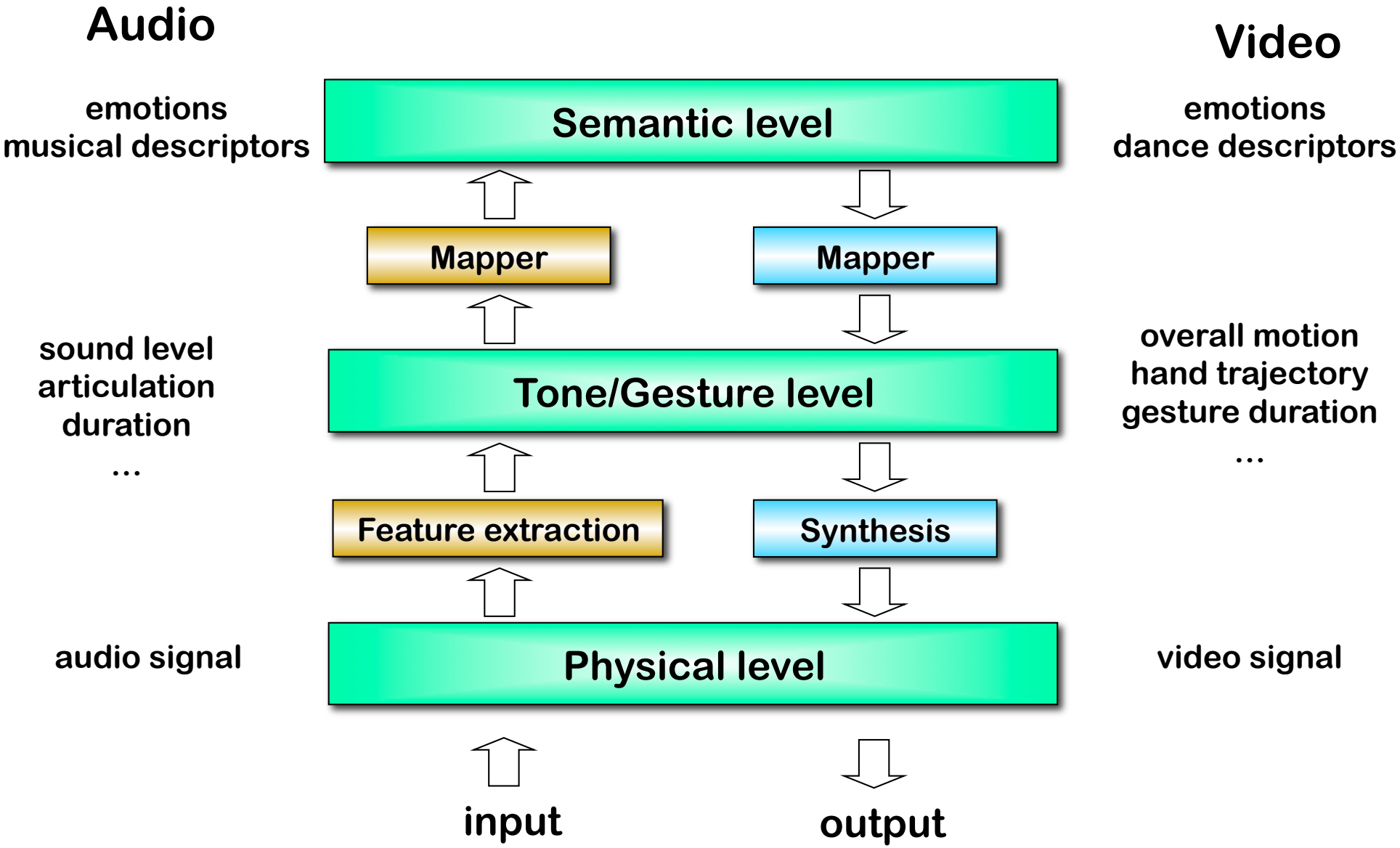
From Leman et al

Framework for analysis and synthesis



From Leman et al

Framework for analysis and synthesis



From Leman et al

Which emotions?

From general emotion research:

Discrete emotions (Ekman…)

Happiness, sadness, anger, tenderness/love, fear

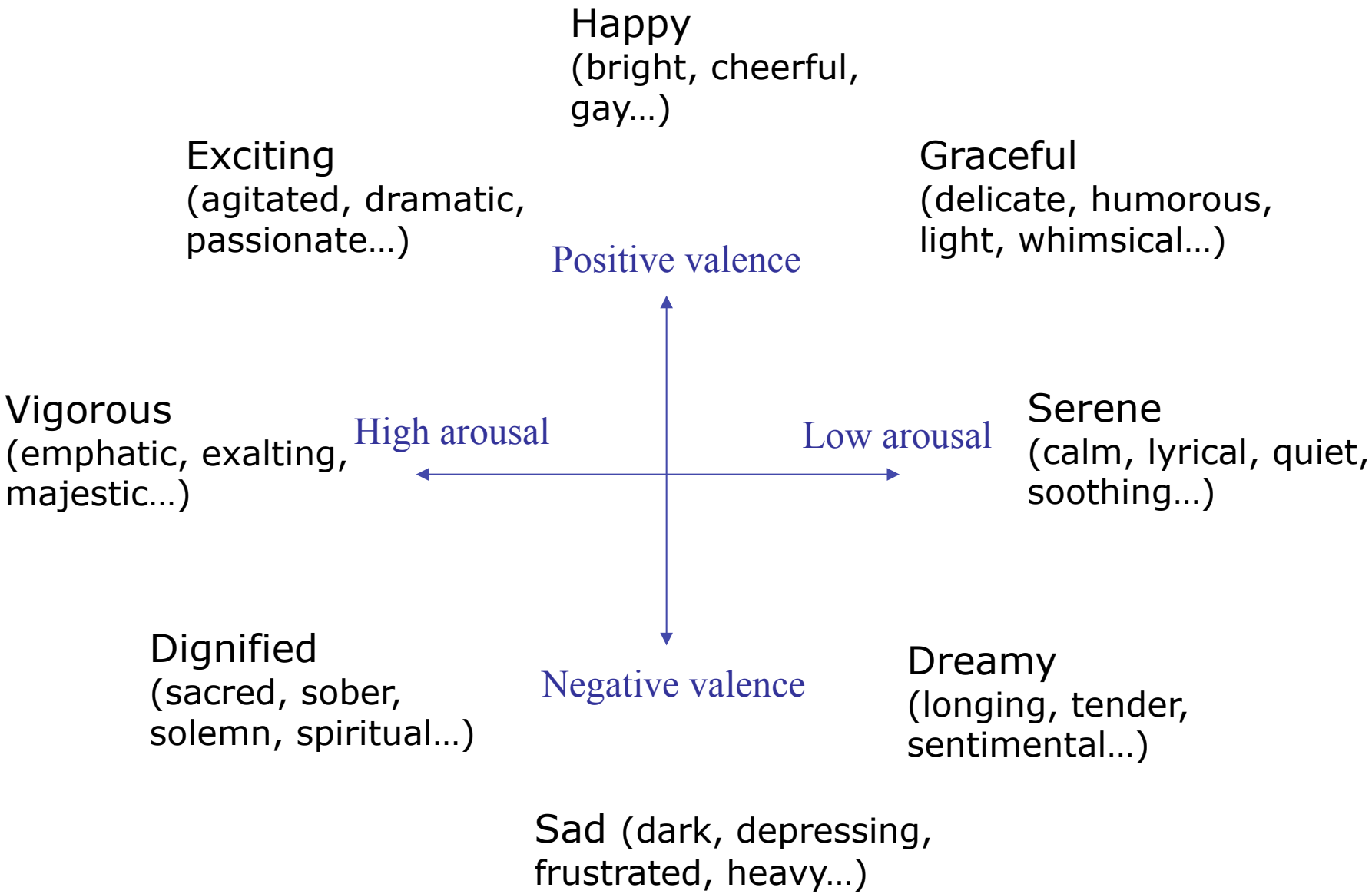
These categories has been used in a number of music studies including the meta-analysis by Juslin and Laukka

Dimensional approach (Russell …)

The most common is a **two dimensional space** with **activity** as one dimension and **valence** as the other dimension. Activity is the associated energy and valence is the positive or negative connotation of the emotion.

Russell (1980) showed that many discrete emotions will be positioned at specific points in the space forming a circle.

Are music emotions special?
Hevner's adjective circle:



Which features?

Performance features

Timing - Tempo, tempo variation, duration contrast
Dynamics: overall level, crescendo/decrescendo, accents
Articulation: overall (staccato/legato), variability
Timbre: Spectral richness, onset velocity

Score features (from Gabrielsson and Lindström 2001)

Pitch (high/low)
Interval (small/large)
Melody range (small/large), direction (up/down)
Harmony (consonant/complex-dissonant)
Tonality (chromatic-atonal/key-oriented)
Rhythm (regular-smooth/firm/flowing-fluent/irregular-rough)
Timbre (harmonic richness)

Features

The set of score features is different from traditional music analysis

For example, instead of the specific harmonic progression a measure of harmonic complexity is more relevant

→ There is a need to develop new features and test them perceptually

→ Good news is that the extraction may be easier than for example traditional phrase analysis or harmonic analysis

Feature extraction

MIDI data

Simpler than audio but not trivial

Existing tools mostly developed for other purposes: Beat extraction, metric analysis, harmonic analysis…

MIDIToolbox, Jyväskylä University (monophonic)

D. Temperley tools

…

Audio data

Substantial development of audio features for other purposes within MIR:

Beat analysis, harmony, mode, main melody, spectral measures …

Beginning of specialized features for emotion:

Pulse clarity, … (MIRToolbox, Jyväskylä U)

Rhythm strength, Rhythm regularity (Lu et al., 2006)

…

Mapping features to emotions

A number of standard methods for data analysis and prediction can be used:

Multiple regression (used in many psychology studies)

Fuzzy logic

Bayesian modelling

Neural networks

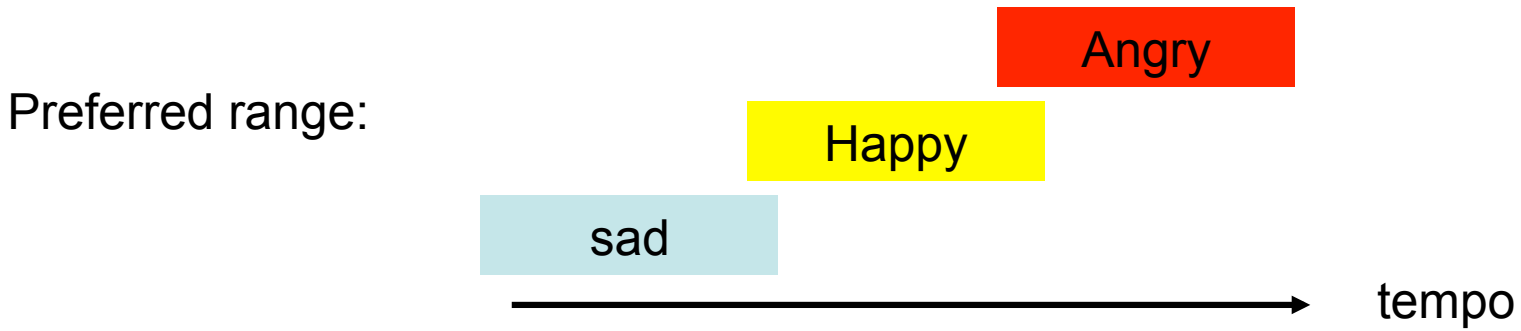
Support Vector Machines (SVM)

…

Multiple regression

+ extremely well investigated and accepted
+ easy to quantize the contribution of each feature
+ not classification

- Its linear behaviour can in some cases cause wrong predictions:



A multiple regression predicting happiness will wrongly conclude that tempo is not important in this case

Analysis of Emotional Expression

Kate Hevner (1935a, 1935b, 1936, 1937) used short excerpts of real music with systematic manipulations

Mode – major/minor

Melodic direction – ascending/descending

Harmony – consonant/dissonant

Rhythm – firm/flowing

Tempo –fast/slow

Pitch level – octave up/down

Played by a pianist to hundreds of subjects that selected adjectives from the circle

Hevner's summary of results from six experiments

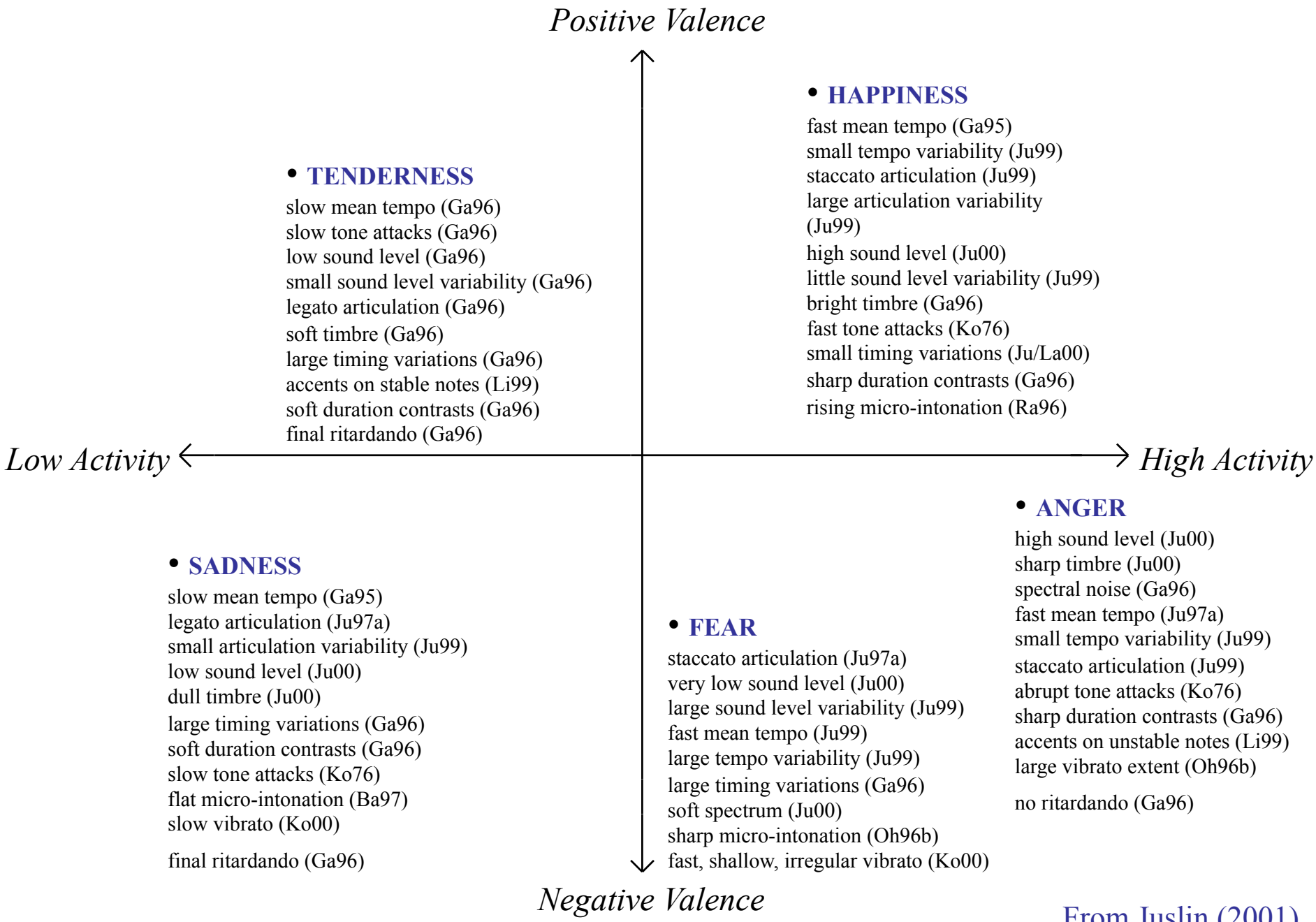
Table 10.1 Hevner's summary of results from six experiments (adapted from Hevner 1937, p. 626)

Musical factor	Dignified/ solemn	Sad/ heavy	Dreamy/ sentimental	Serene/ gentle	Graceful/ sparkling	Happy/ bright	Exciting/ elated	Vigorous/ majestic
Mode	Minor 4	Minor 20	Minor 12	Major 3	Major 21	Major 24	–	–
Tempo	Slow 14	Slow 12	Slow 16	Slow 20	Fast 6	Fast 20	Fast 21	Fast 6
Pitch	Low 10	Low 19	High 6	High 8	High 16	High 6	Low 9	Low 13
Rhythm	Firm 18	Firm 3	Flowing 9	Flowing 2	Flowing 8	Flowing 10	Firm 2	Firm 10
Harmony	Simple 3	Complex 7	Simple 4	Simple 10	Simple 12	Simple 16	Complex 14	Complex 8
Melody	Ascend 4	–	–	Ascend 3	Descend 3	–	Descend 7	Descend 8

The numbers indicate the relative weight of each musical factor (left column) for each emotion cluster.

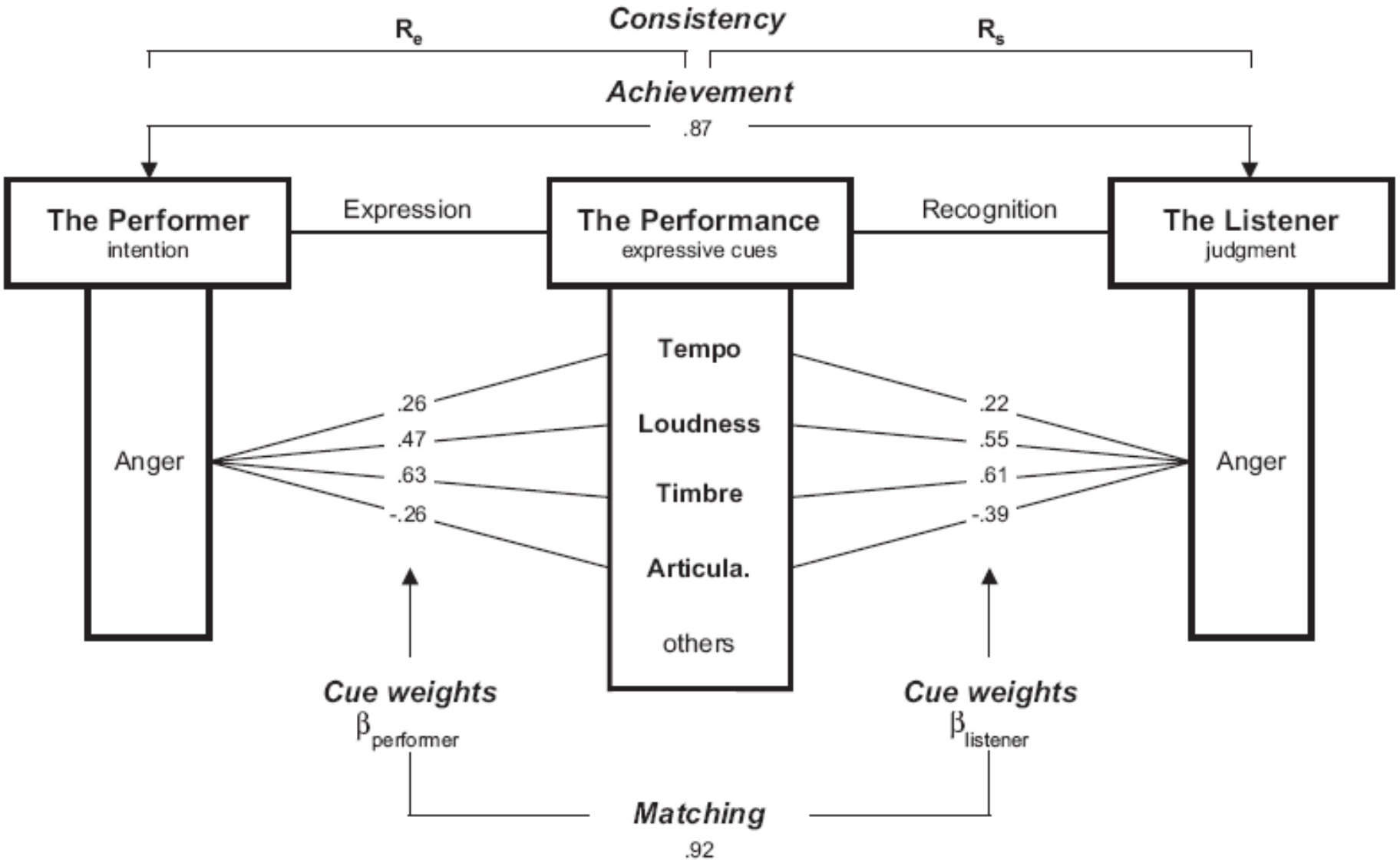
From Gabrielsson and Lindström 2001

Qualitative estimates of performance features



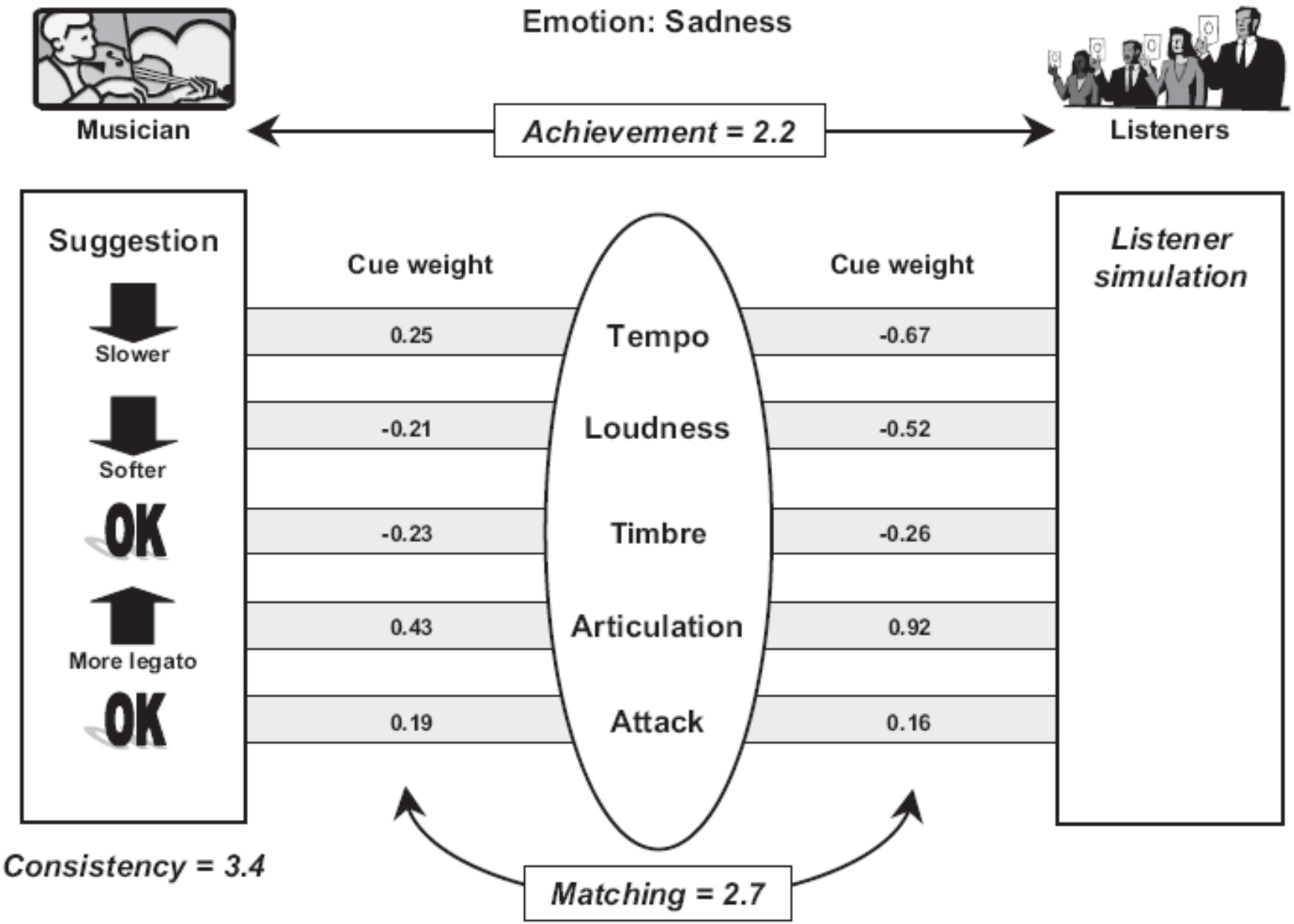
From Juslin (2001)

Juslin's lens model: quantifies the expressive communication between performer and listener



From Juslin 2000

Feel-Me project: learning to express emotions by automatic cognitive feedback



From Juslin et al. 2004,2006

Recent MIR oriented studies

Leman et al. 2005, JNMR

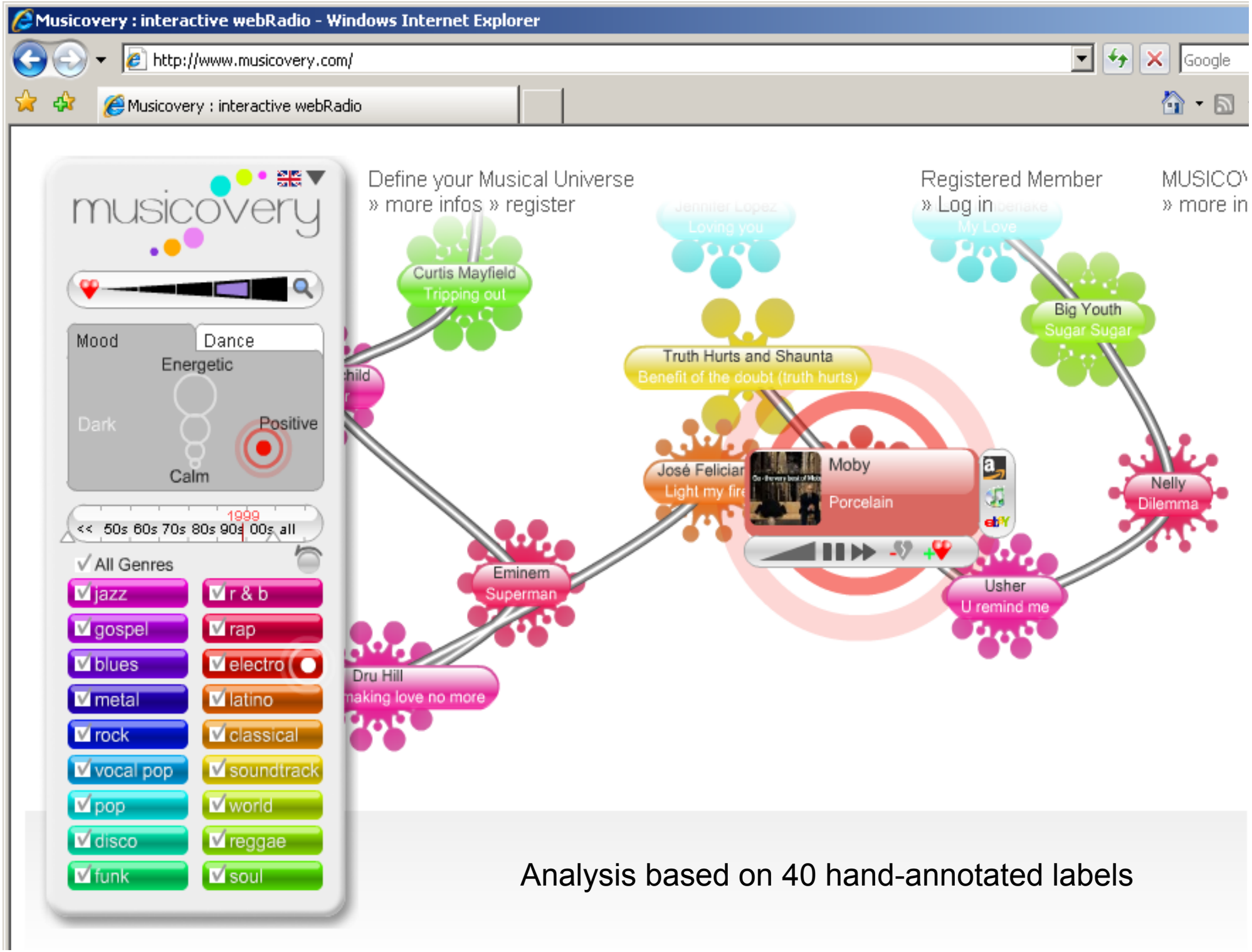
Large number of features, both automatic and annotated by music experts

Large number of emotion labels

60 music examples, different styles

Factor analysis of 100 subjects → three main dimensions: valence activity and interest

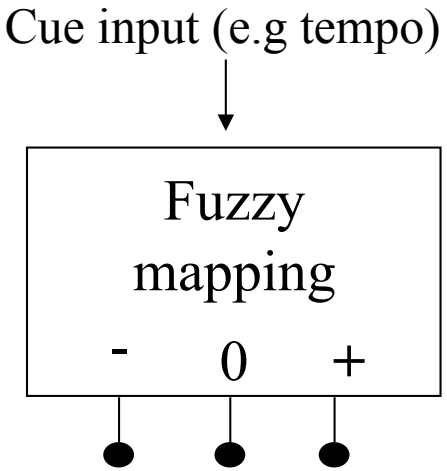
Regression analysis predicting the dimensions: activity ok, valence somewhat and interest not



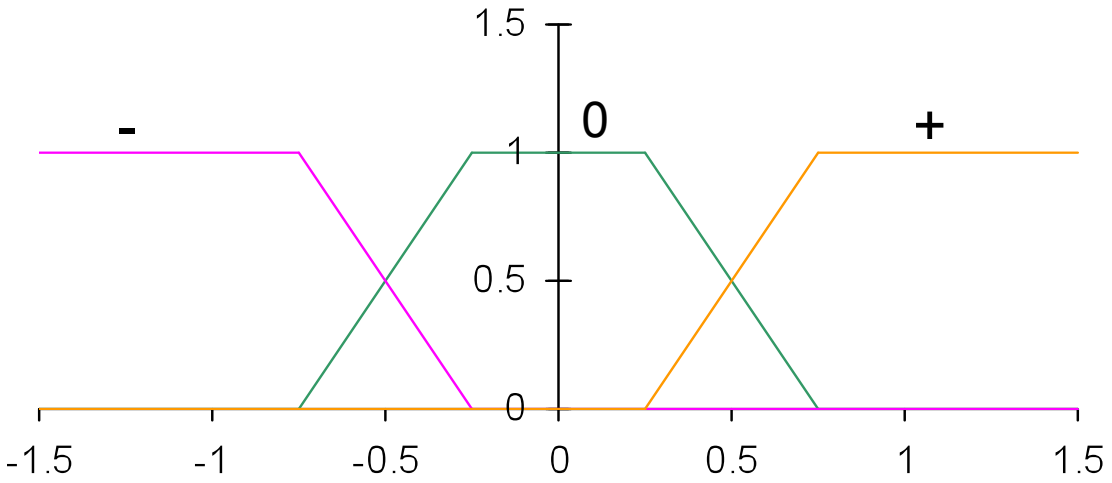
Analysis based on 40 hand-annotated labels

Simple real-time analysis

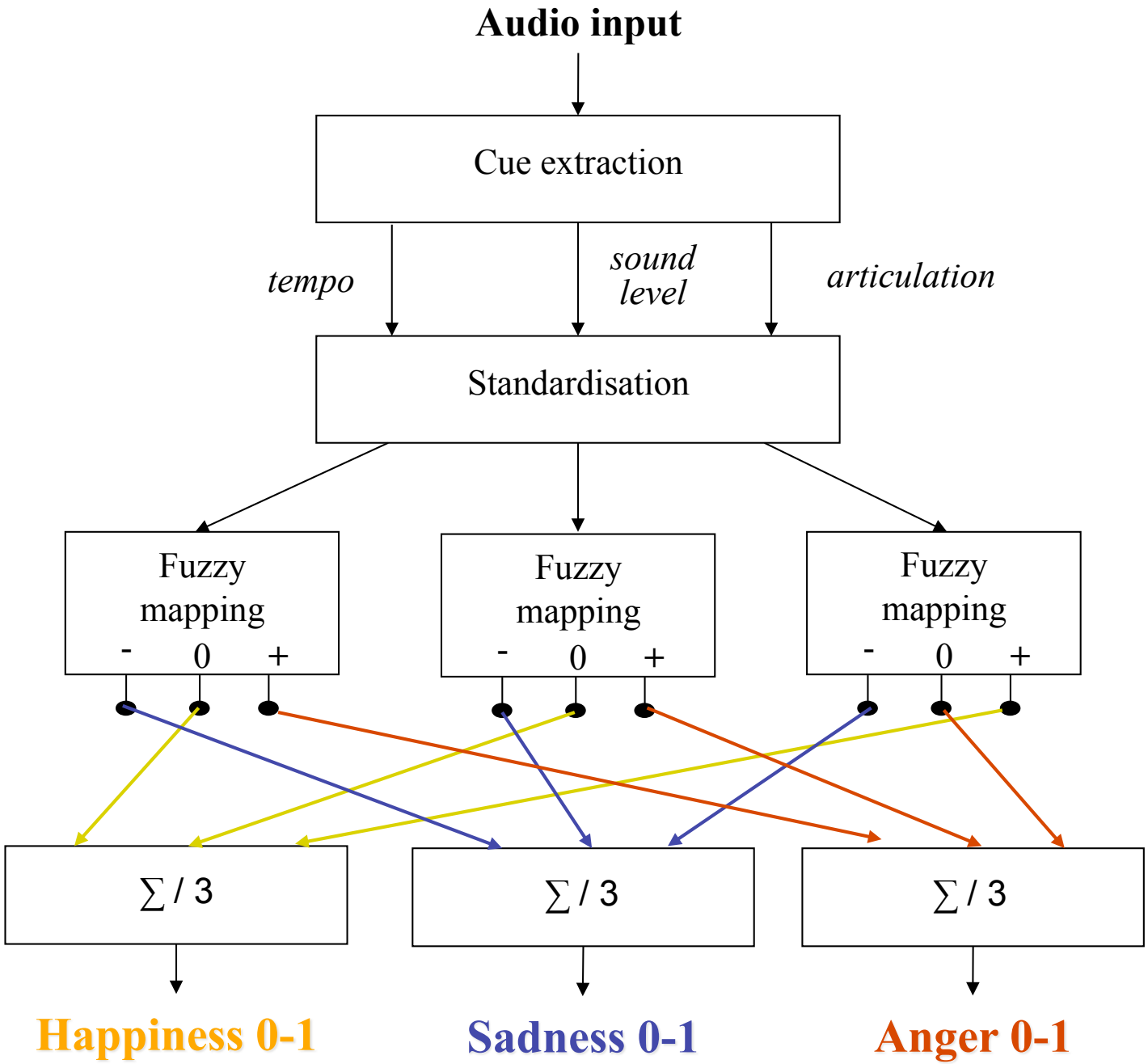
FUZZY MAPPING
from cue variables to emotion recognition



"Qualitative Classification"



From audio to emotion recognition



Synthesis of Emotional Expression

Mapping from Emotional Expression to KTH Rule System for Musical Expression

For each emotion:

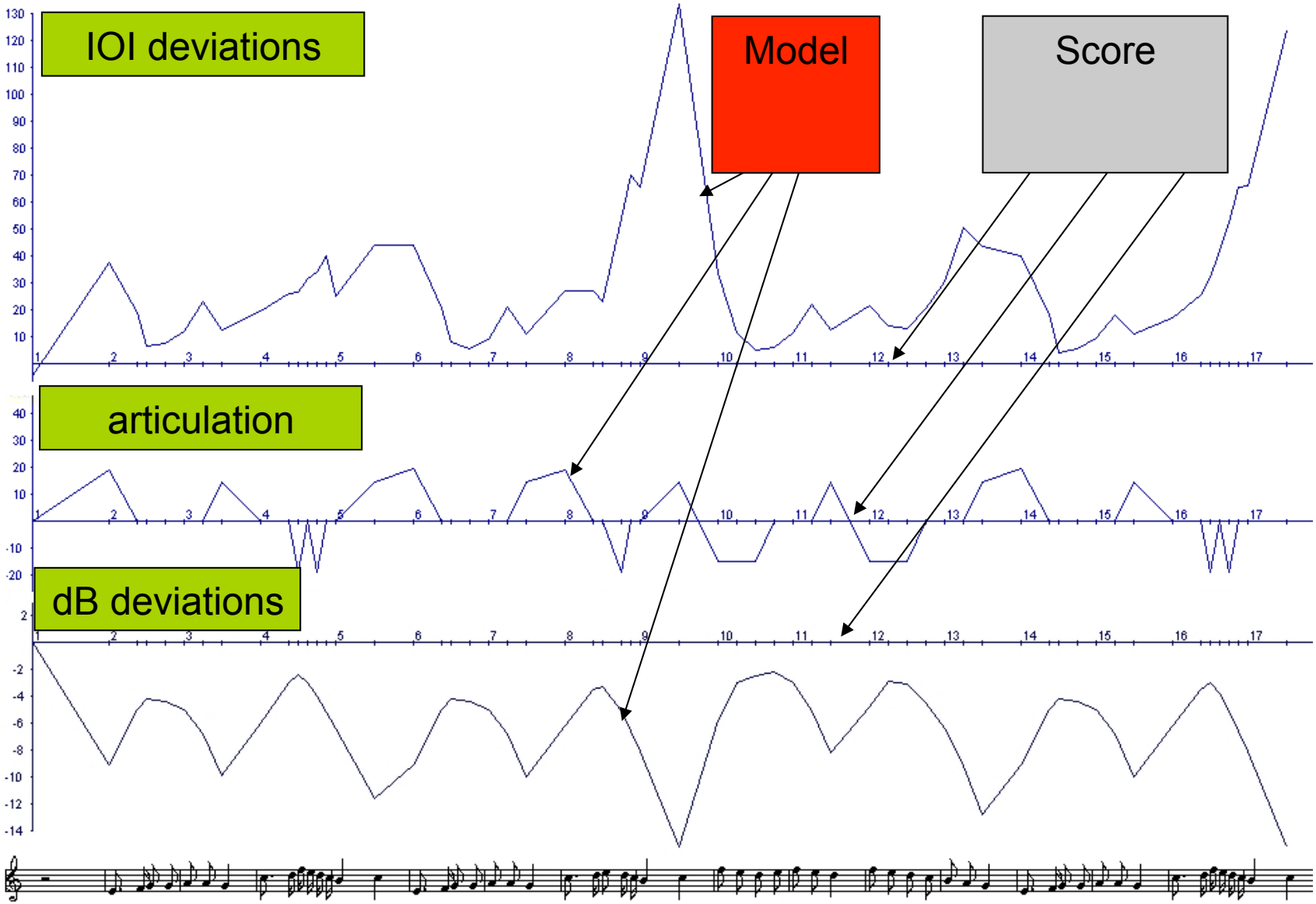
Select a palette of rule parameters according to previous findings

Synthesis of SADNESS

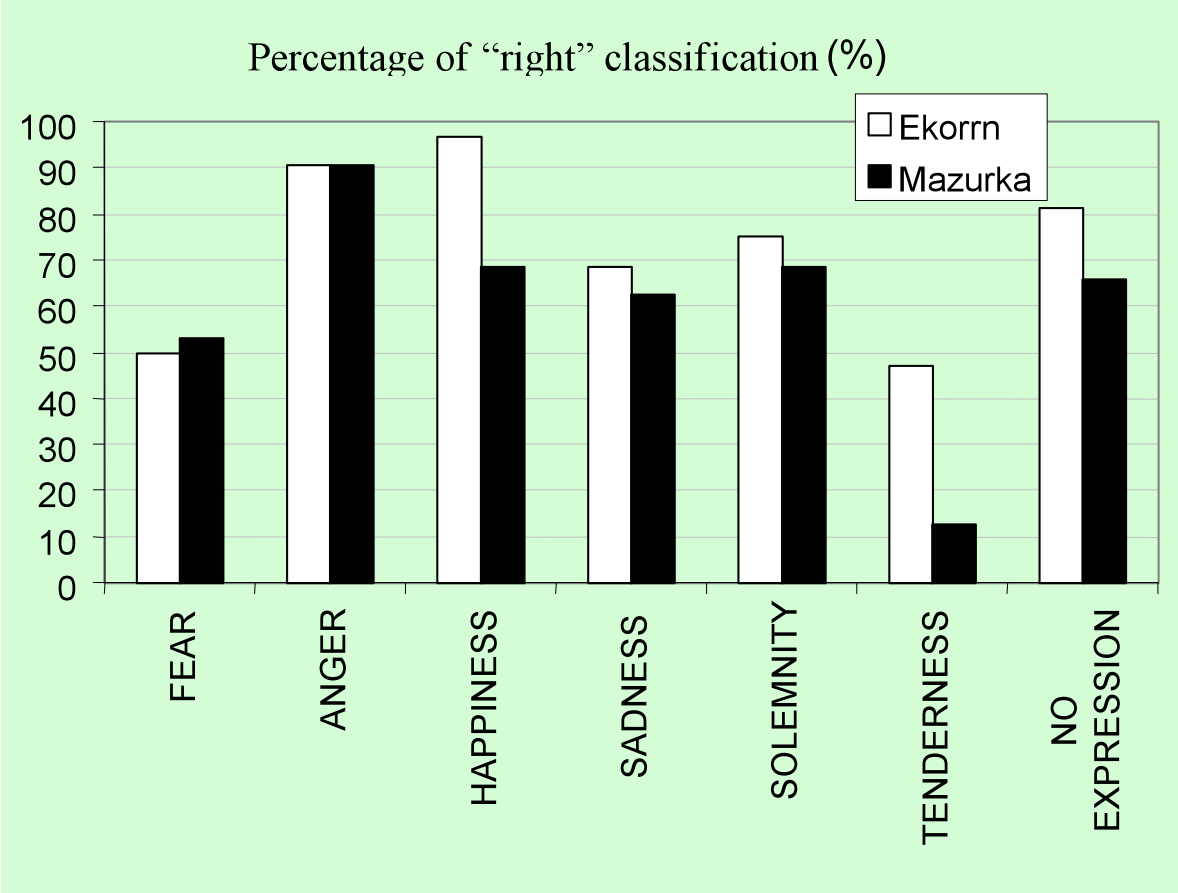
Expressive Cue	Analysis	Synthesis (Director Musices)
Tempo	Slow	Tone IOI is lengthened by 30%
Sound level	Moderate or low	Sound level is decreased by 6 dB
Articulation	Legato	Tone duration = Tone IOI
Time deviations	Moderate	♣ Duration Contrast Rule (k = -2) ♣ Phrase Arch Rule applied on phrase level (k = 1.5) ♣ Phrase Arch Rule applied on sub-phrase level (k = 1.5)
Final ritardando	Yes	Obtained from the Phrase Rule

Bresin and Friberg 2000

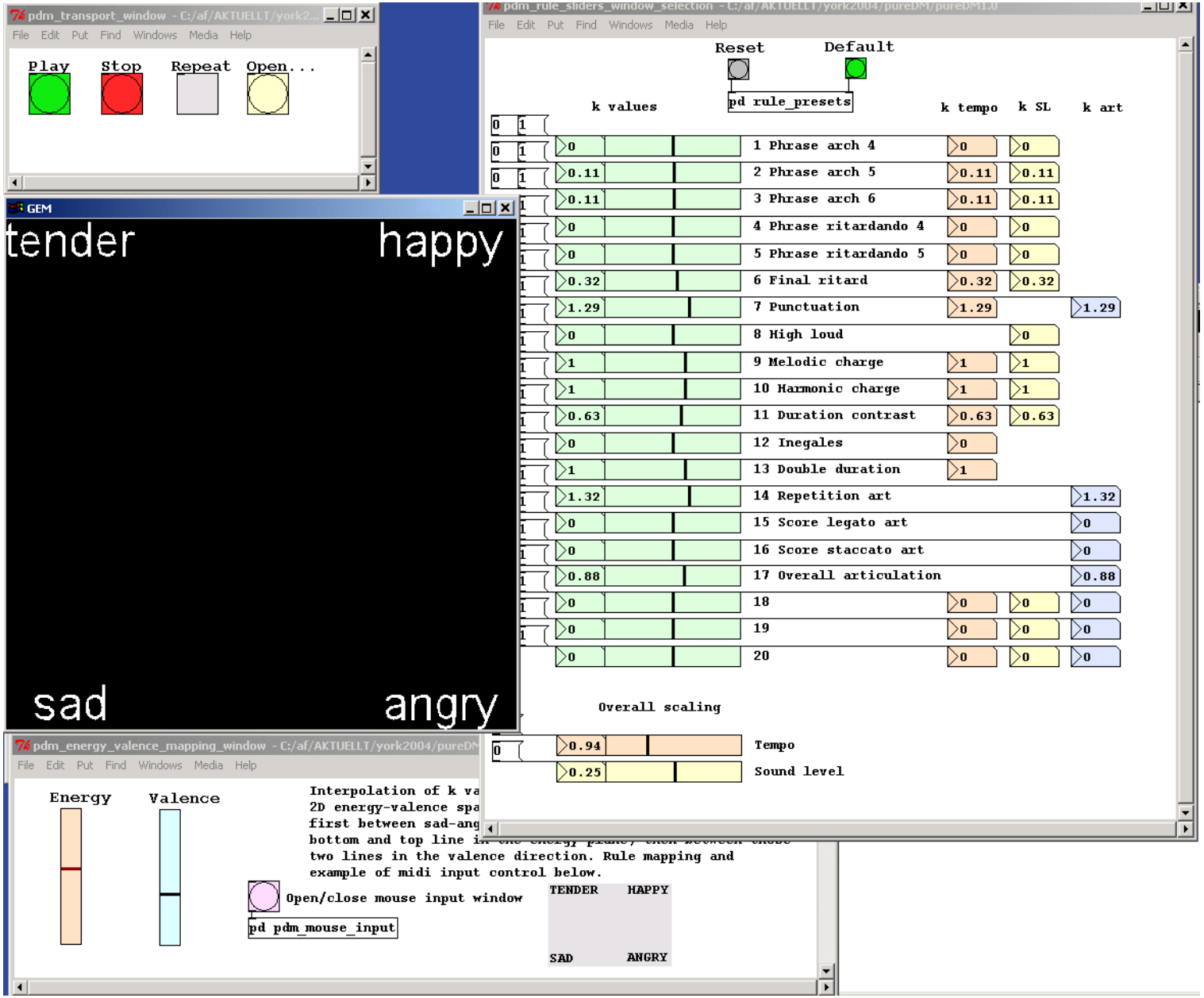
Example: SADNESS



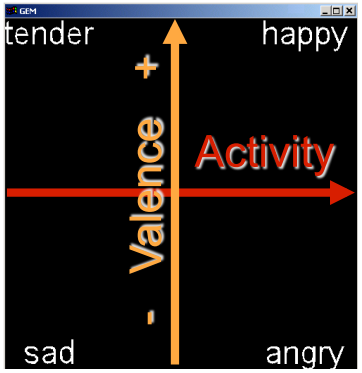
Synthesis of Emotion: Listening Test Results

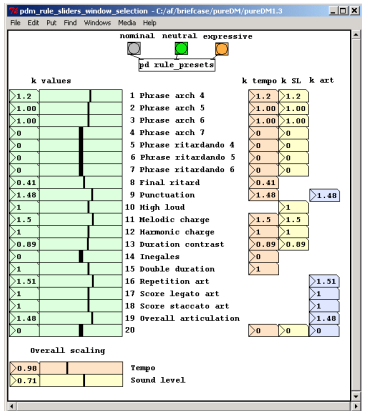


pDM –performance rules in real-time



Complete Home conductor system





Expressive gestures

Gesture cue extraction

Motion data

Mapper

High-Level Expression

Mapper

Rule parameters

pDM ← *Score*

Tone instructions

MIDI synthesizer

Sound