The role of acoustic and perceptual features in *music* perception

Pauline Larrouy-Maestri,[1,2] T. Ata Aydin,[1] & Melanie Wald-Fuhrmann[1]
[1] Department of Music, Max Planck Institute for Empirical Aesthetics, Frankfurt am Main, Germany
[2] Center for Language, Music, and Emotion (CLaME), New York, United States

Background

- Long standing question: "What is music?"
- Previous findings: (Larrouy-Maestri & Wald-Fuhrmann, pre-registration)
  - People identify sounds as music or not confidently.
  - The perception of music is slightly affected by
    - listeners' perspective (Exp. 1)
    - stimuli duration (Exp. 2)
    - stimuli repetition (Exp. 3)

    BUT these effects are small.
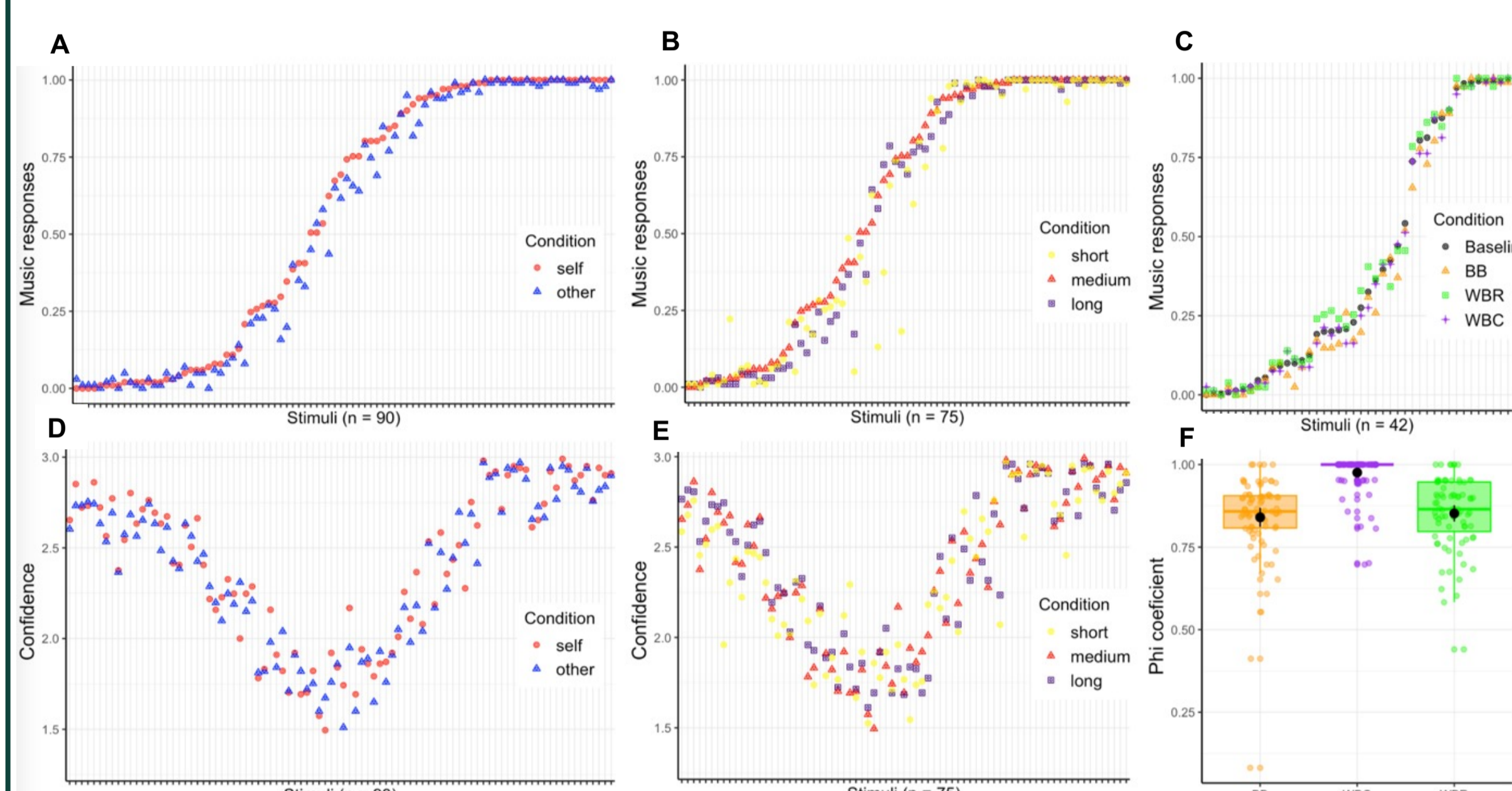
Mean music answers (A, B, C, yes/no answers) and confidence ratings (D, E, ordinal scale: 0-3, from not at all to very confident) for each stimulus and each condition of the pre-registered experiments testing the effect of listeners perspective (Exp. 1), stimuli's duration (Exp. 2) and repetition (Exp. 3), on the participants (n = 637) identification of audio stimuli (n = 90, 75, 42, respectively). F illustrates listeners' consistency when identifying the 42 stimuli two times, estimated with Phi-coefficients between first and second presentations (i.e., Between Blocks Random, Within Blocks Consecutive).

This experiment

Role of <u>acoustic</u> features in *music* perception?

Evidence FOR and AGAINST a music-acoustic mapping
- Numerous studies comparing speech and song (e.g., Albouy et al., 2024; Bruder et al., preprint; Chang et al., 2024; Ozaki et al., 2024)
- Speech-to-Song illusion (Deutsch et al., 2011 and many others)
- Music is extremely varied, within/between cultures, and over time (e.g., Titon, 2016)
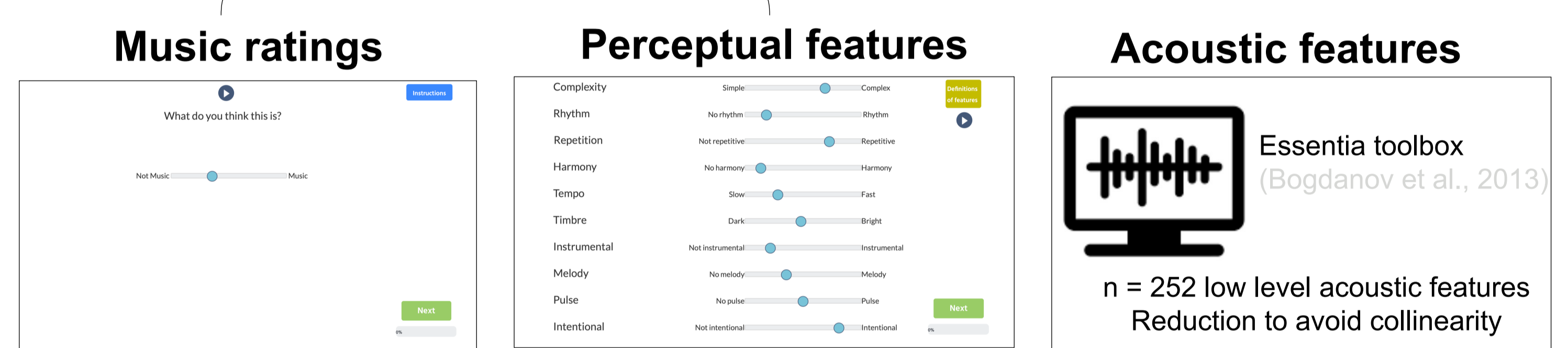
Role of <u>perceptual</u> features in *music* perception?

In the context of singing voices, perceptual features (vibrato, attack, brightness, etc.) predict listeners' preferences (Bruder et al., 2024)

Methods

Material: 90 stimuli from different sources (doi: 10.17617/3.I9BJQ1)

98 online Western participants

Music ratings | **Perceptual features** | **Acoustic features**

Perceptual features:
Complexity, Rhythm, Repetition, Harmony, Tempo, Timbre, Instrumental, Melody, Pulse, Intentional

Music ratings: What do you think this is? Not Music — Music

Acoustic features: Essentia toolbox (Bogdanov et al., 2013)
n = 252 low level acoustic features
Reduction to avoid collinearity

Principal Component Analysis: reduction to two-dimensional spaces

Syntax: Music ratings ~ PC1 + PC2 + (1|Participants) + (1|Stimuli)

Results
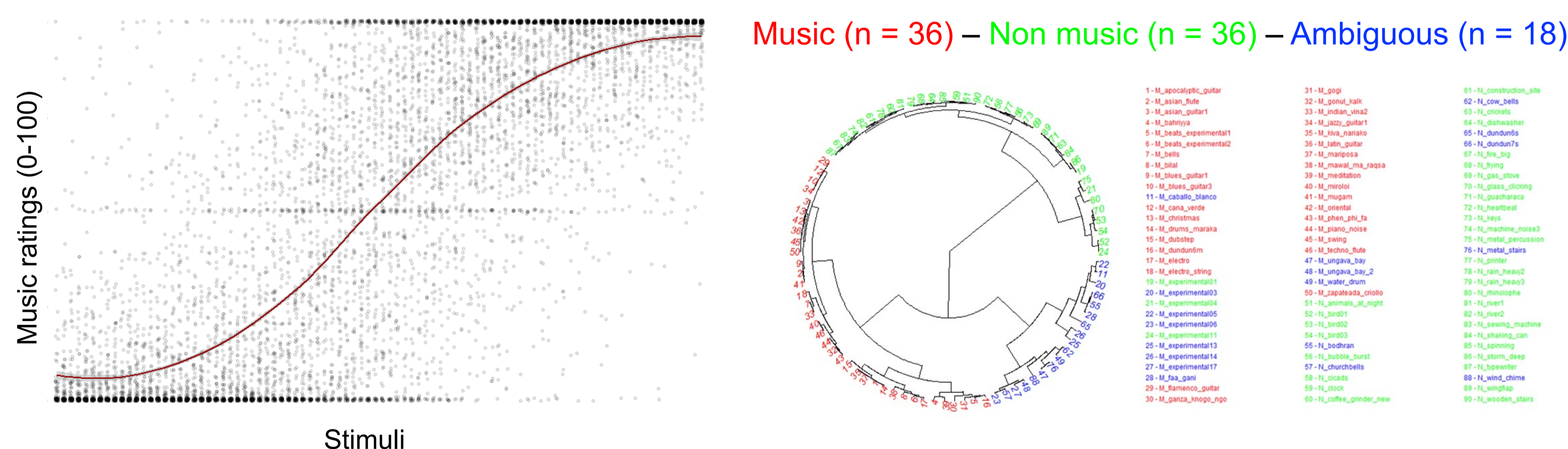
Ratings on the music slider (0-100) of the 90 stimuli by Western online participants (n = 98). Individual points represent the raw ratings and the curve is the fitted sig function of the average ratings.

Music (n = 36) – Non music (n = 36) – Ambiguous (n = 18)

Dendrogram representing the grouping similarity of the 90 stimuli. Small heights indicate greater similarity. Colors represent **clusters** (n = 3, according to the visual inspection of k-means cluster analysis).
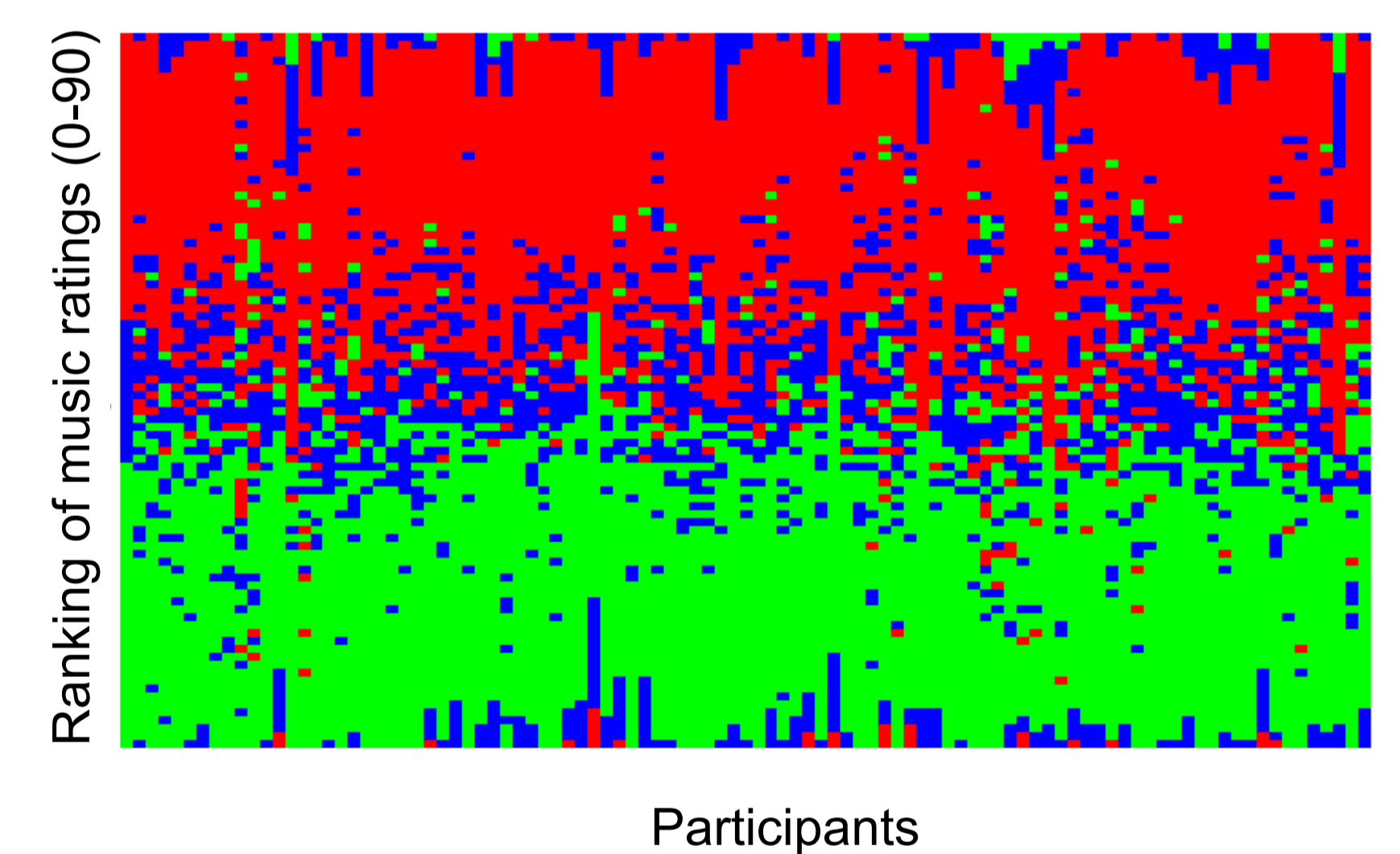
Moderate variability between participants (n = 98 columns). For each participant, stimuli are ranked from 1 to 90 and color-coded according to the cluster analysis performed at the group level.
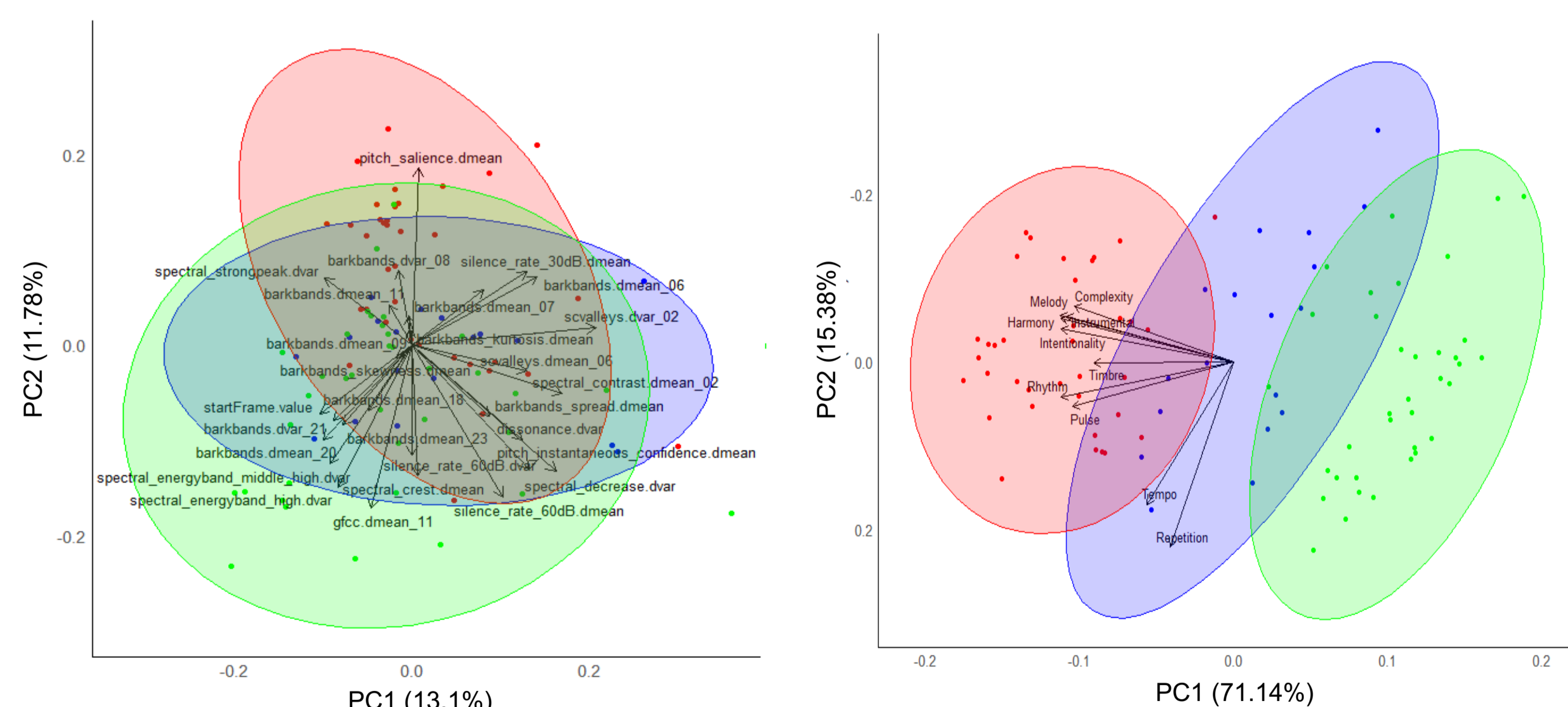
Three clusters on the two-dimensional **acoustic** space. Loadings of the main features on the two dimensions are represented with arrows. Proportion of variance explained by the dimensions: R^2 = 26.5%

Three clusters on the two-dimensional **perceptual** space. Loadings of the features on the two dimensions are represented with arrows. Proportion of variance explained by the dimensions: R^2 = 66.1%
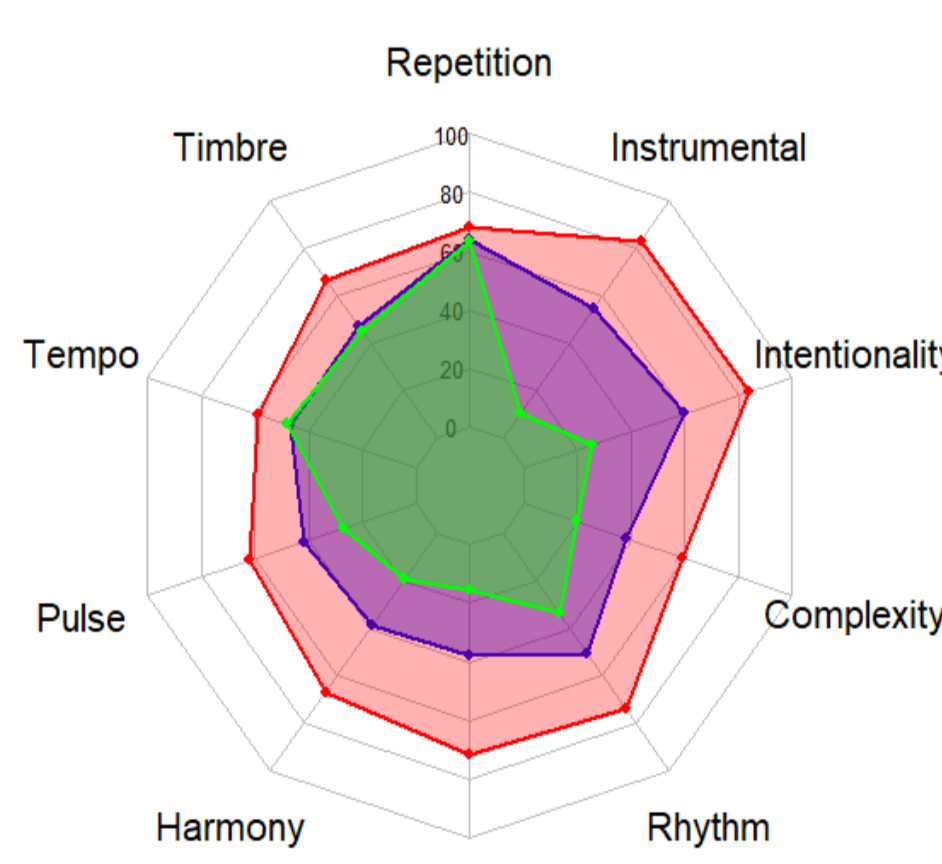
Radar chart highlighting the differences in mean ratings of features between clusters. *Instrumental, Intentionality,* and *Melody* are particularly relevant features, as confirmed by the model including all perceptual features as fixed effects.

Take home messages

- The perception of *music* is stable across conditions (with slight individual differences)
- There are three categories: Music, not music, and ambiguous
- **Perceptual (but not acoustic) features ground *music* perception**

References
Albouy et al. (2024). *Nature Communications;* Bogdanov et al. (2013). *14th International Society for MIR Conference;* Bruder et al. (2024). *Scientific Reports;* Bruder et al. (under review); Deutsch et al. (2011). *JASA;* Chang et al. (2024). Plos Biology; Larrouy-Maestri & Wald-Fuhrmann (2022-2023). OSF: osf.io/sb24q, osf.io/azkcp, osf.io/re7dg; Ozaki et al. (2024). *Science Advances;* Titon (2016). Worlds of Music