Validating a technique for post-hoc estimation of a listener's focus in music structure analysis

Jordan B. L. Smith

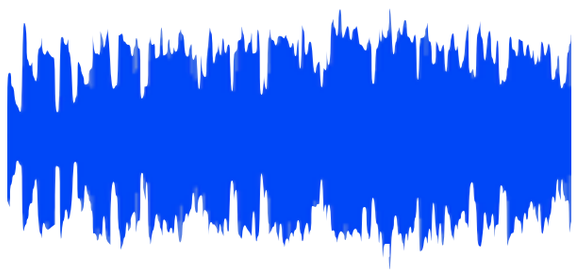
National Institute of Advanced Industrial Science and Technology (AIST), Japan

Elaine Chew

Queen Mary University of London

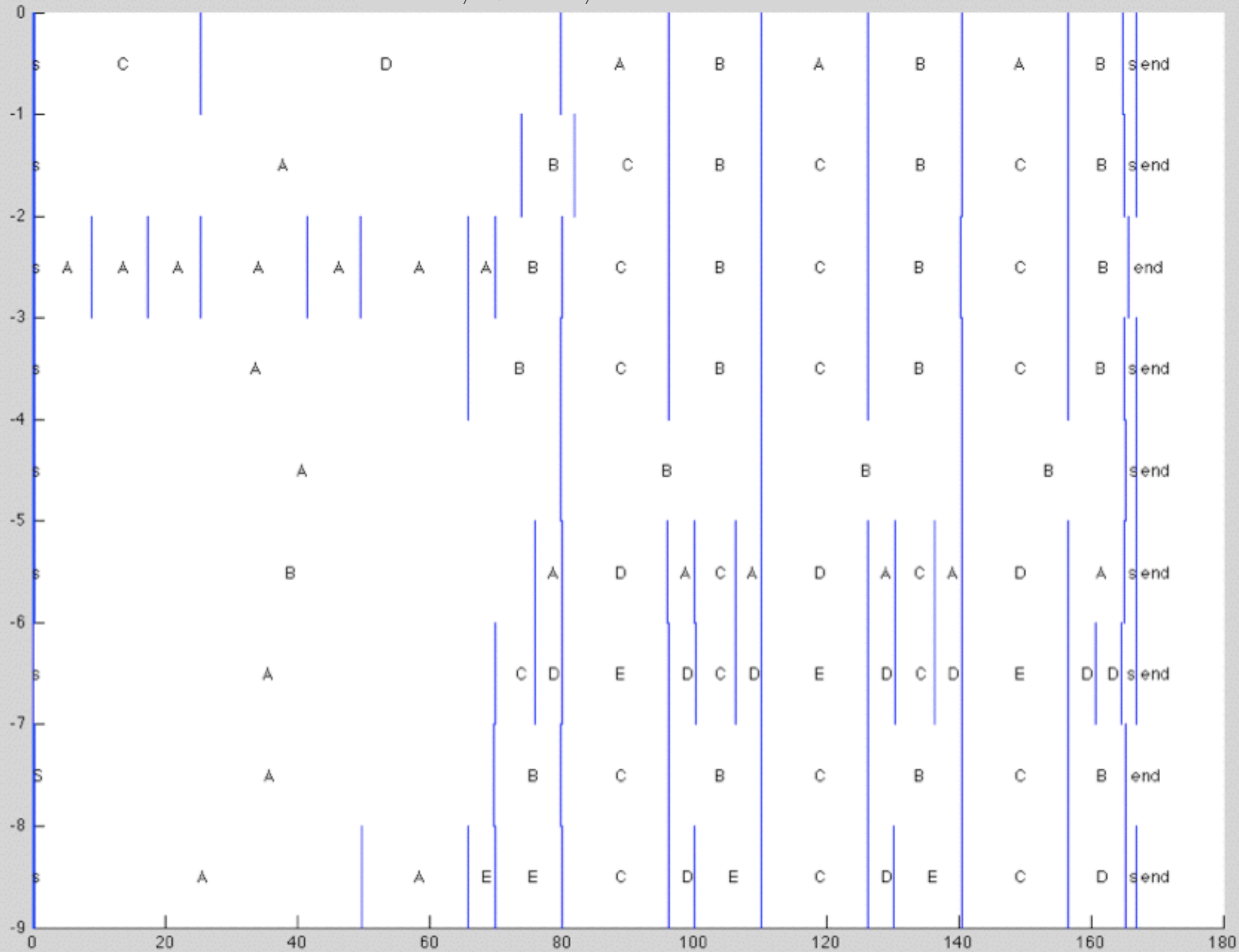
What was the listener thinking?

(when they annotated musical structure)

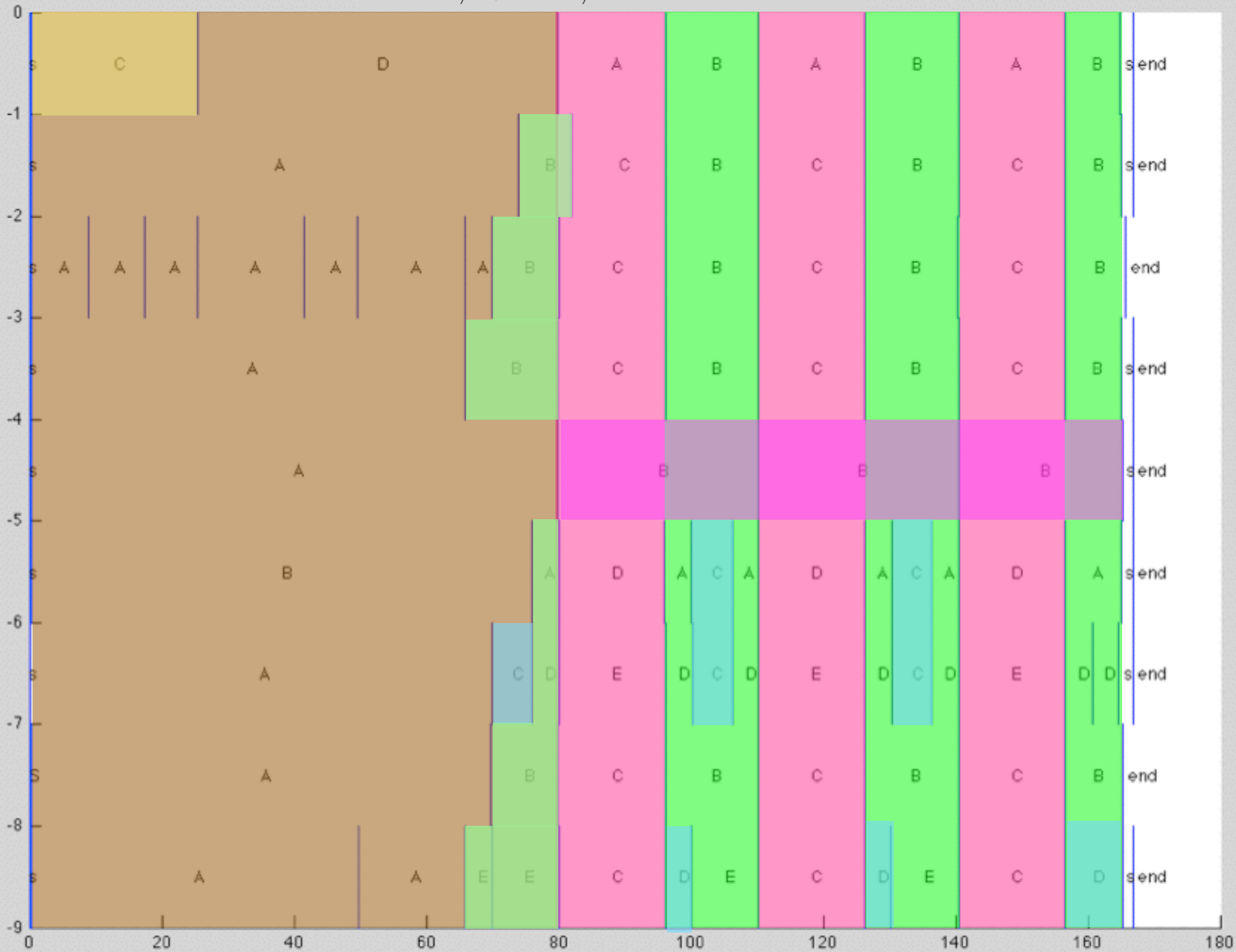


Melody
Harmony
Rhythm
Meter
Timbre
etc.

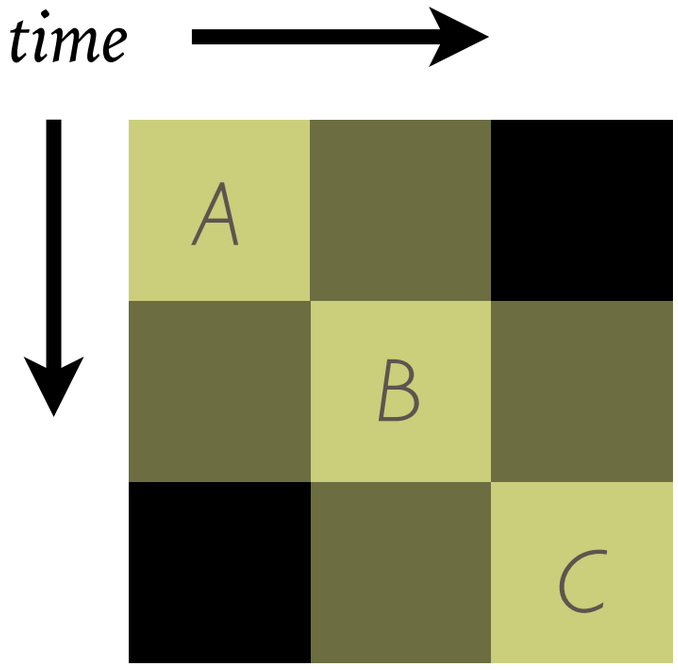
Isaac Hayes, Run Fay Run



Isaac Hayes, Run Fay Run

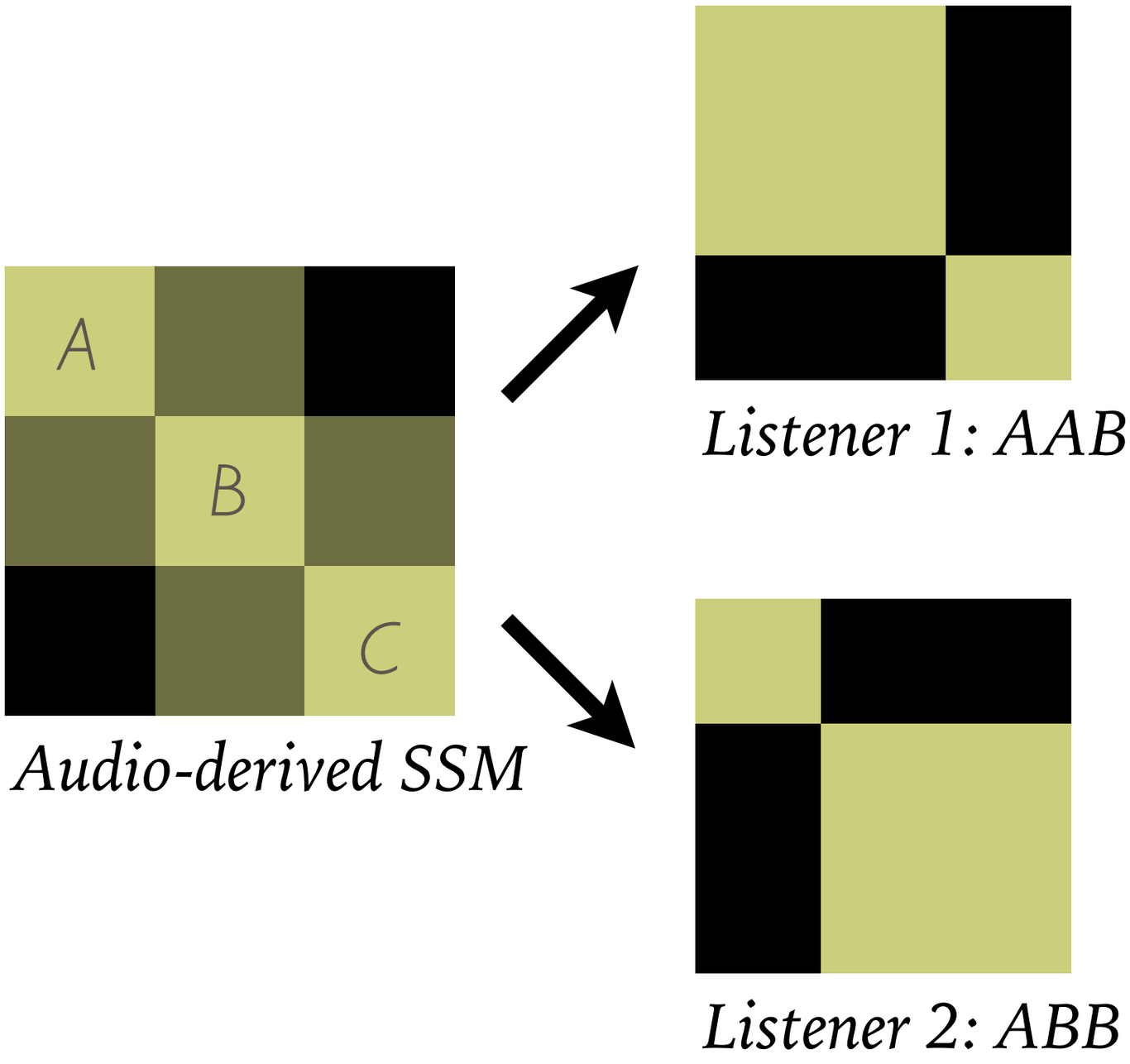


RELATING GROUPING STRUCTURE TO MUSICAL FEATURES

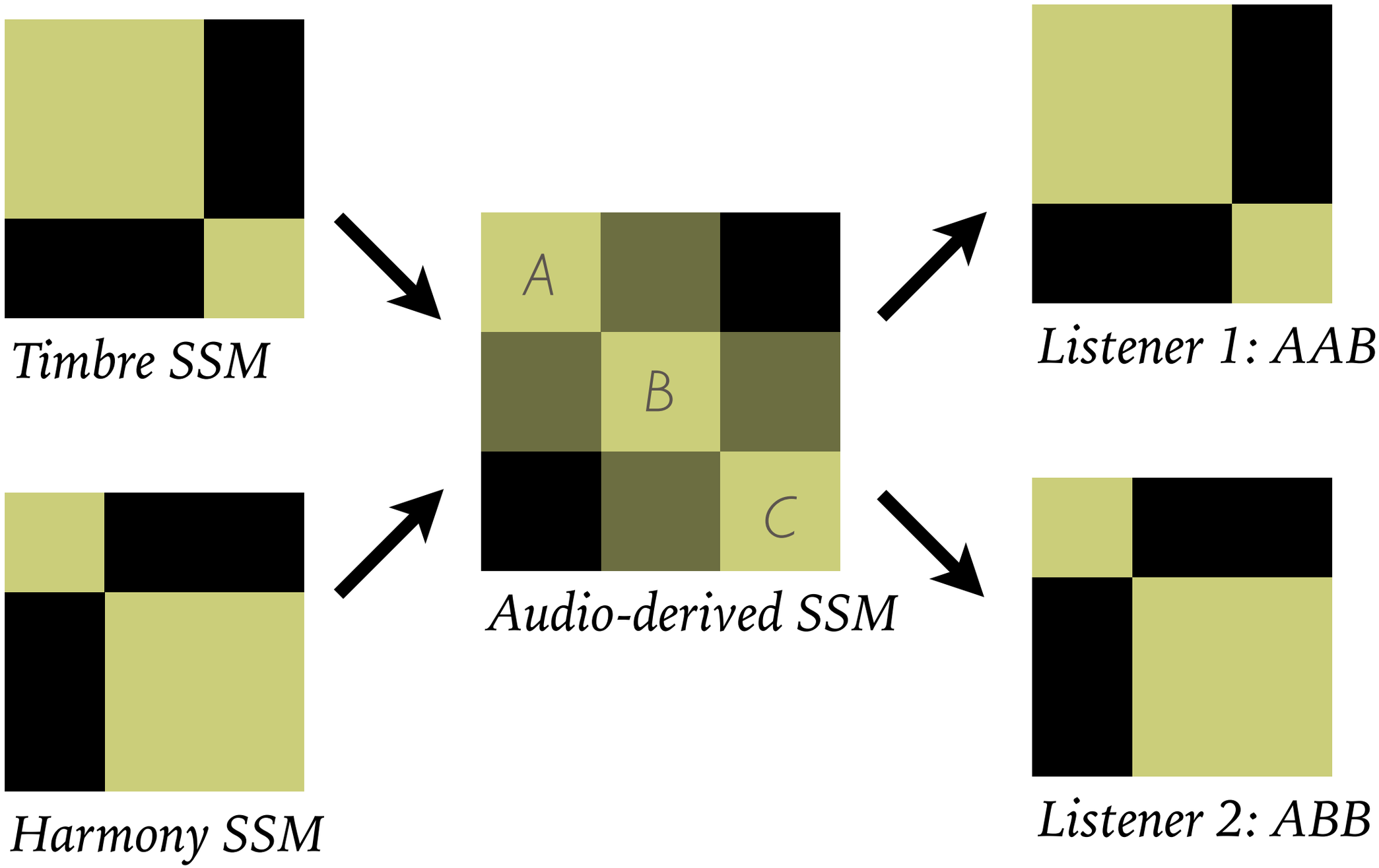


Audio-derived SSM

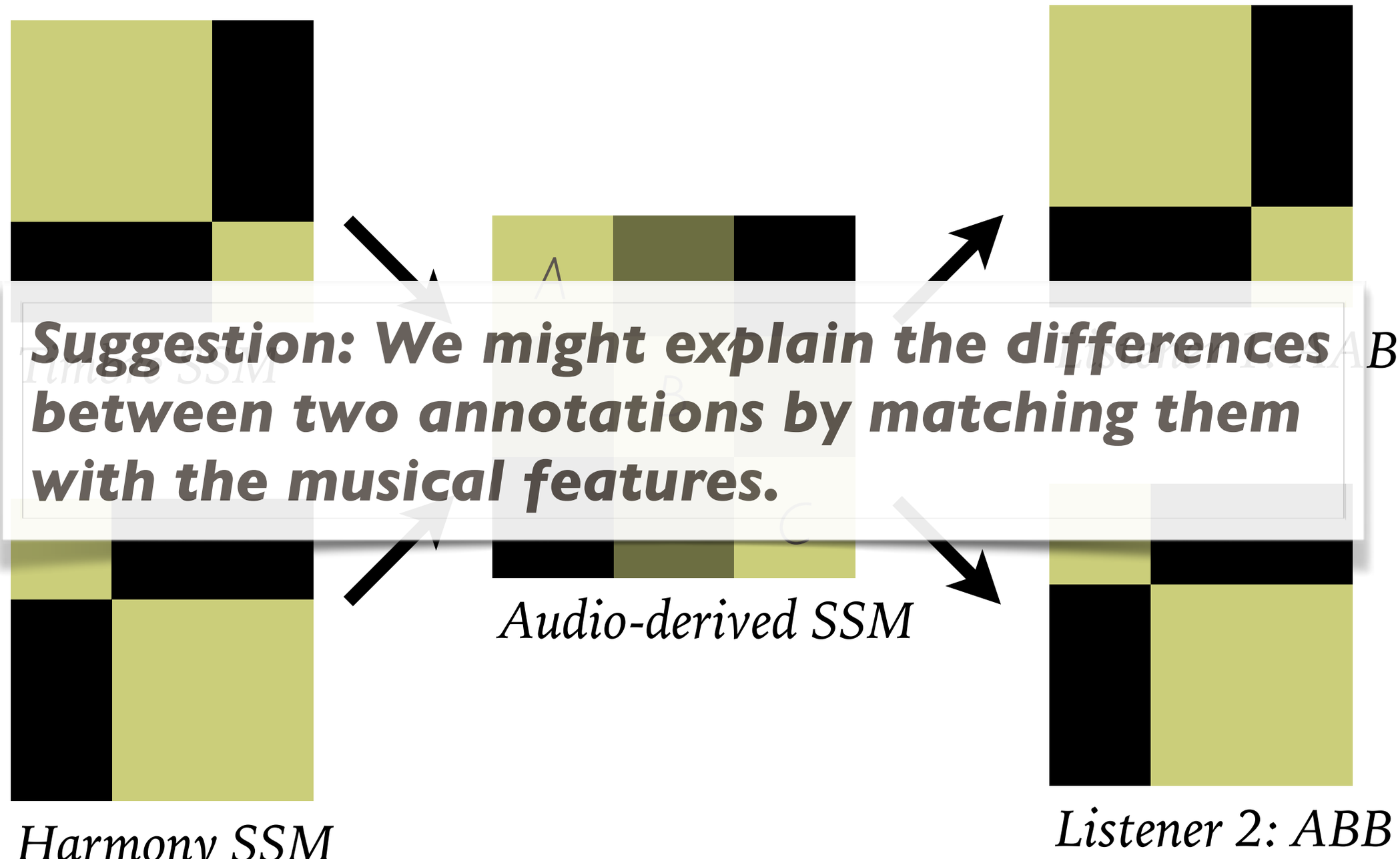
RELATING GROUPING STRUCTURE TO MUSICAL FEATURES



RELATING GROUPING STRUCTURE TO MUSICAL FEATURES



RELATING GROUPING STRUCTURE TO MUSICAL FEATURES

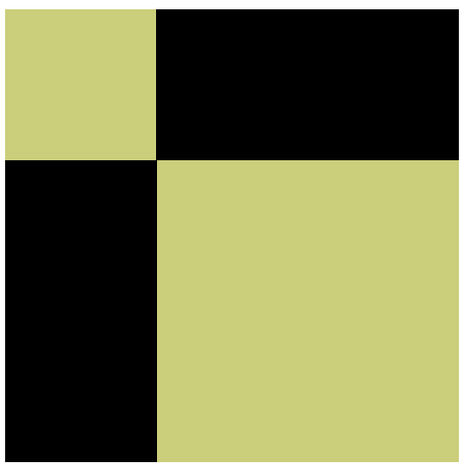


Suggestion: We might explain the differences between two annotations by matching them with the musical features.

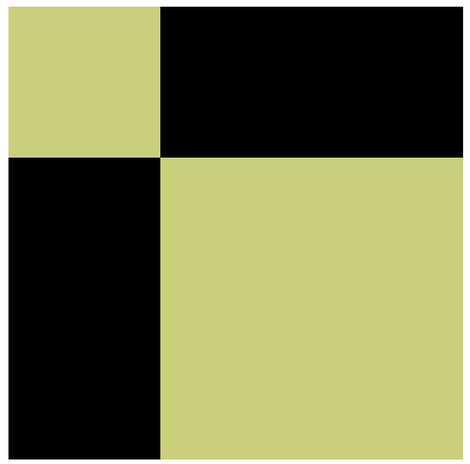
RELATING GROUPING STRUCTURE TO MUSICAL FEATURES

Timbre SSM

Listener 1: AAB



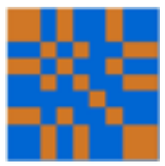
Harmony SSM



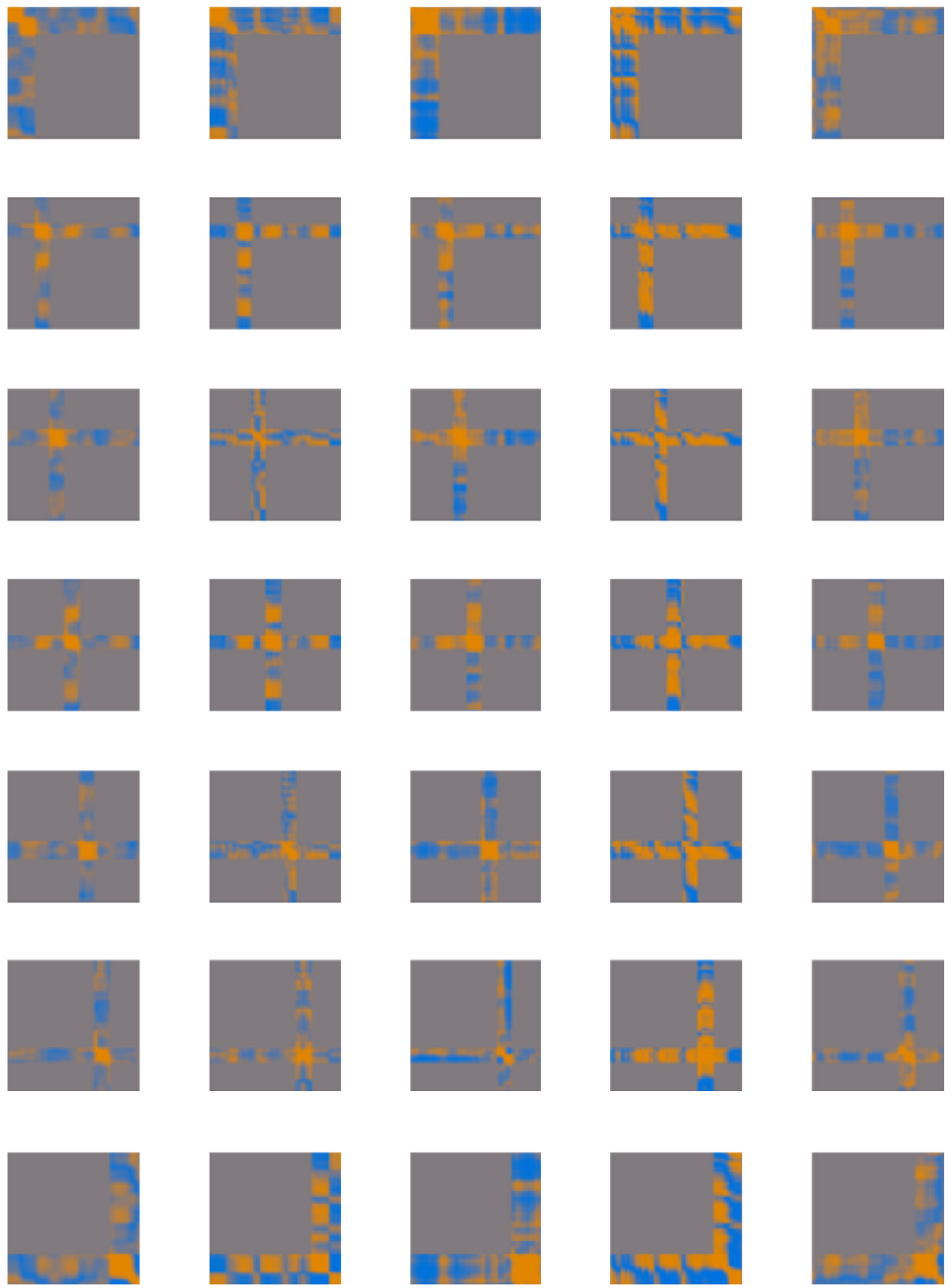
Listener 2: ABB

Target SSM (original annotation)

Feature-derived SSMs



Segment-based masks



Reconstruction components

Target SSM (original annotation)

Feature-derived SSMs

N

$x_{11}M_{11}$ $x_{12}M_{12}$ $x_{13}M_{13}$ ... $x_{15}M_{15}$

$x_{21}M_{21}$

$x_{31}M_{31}$

Segment-based masks

$x_{71}M_{71}$ $x_{75}M_{75}$

Reconstruction components

Target SSM (original annotation)

Segment-based masks

Feature-derived SSMs

N

$x_{11}M_{11}$ $x_{12}M_{12}$ $x_{13}M_{13}$ ... $x_{15}M_{15}$

$x_{21}M_{21}$

$x_{31}M_{31}$

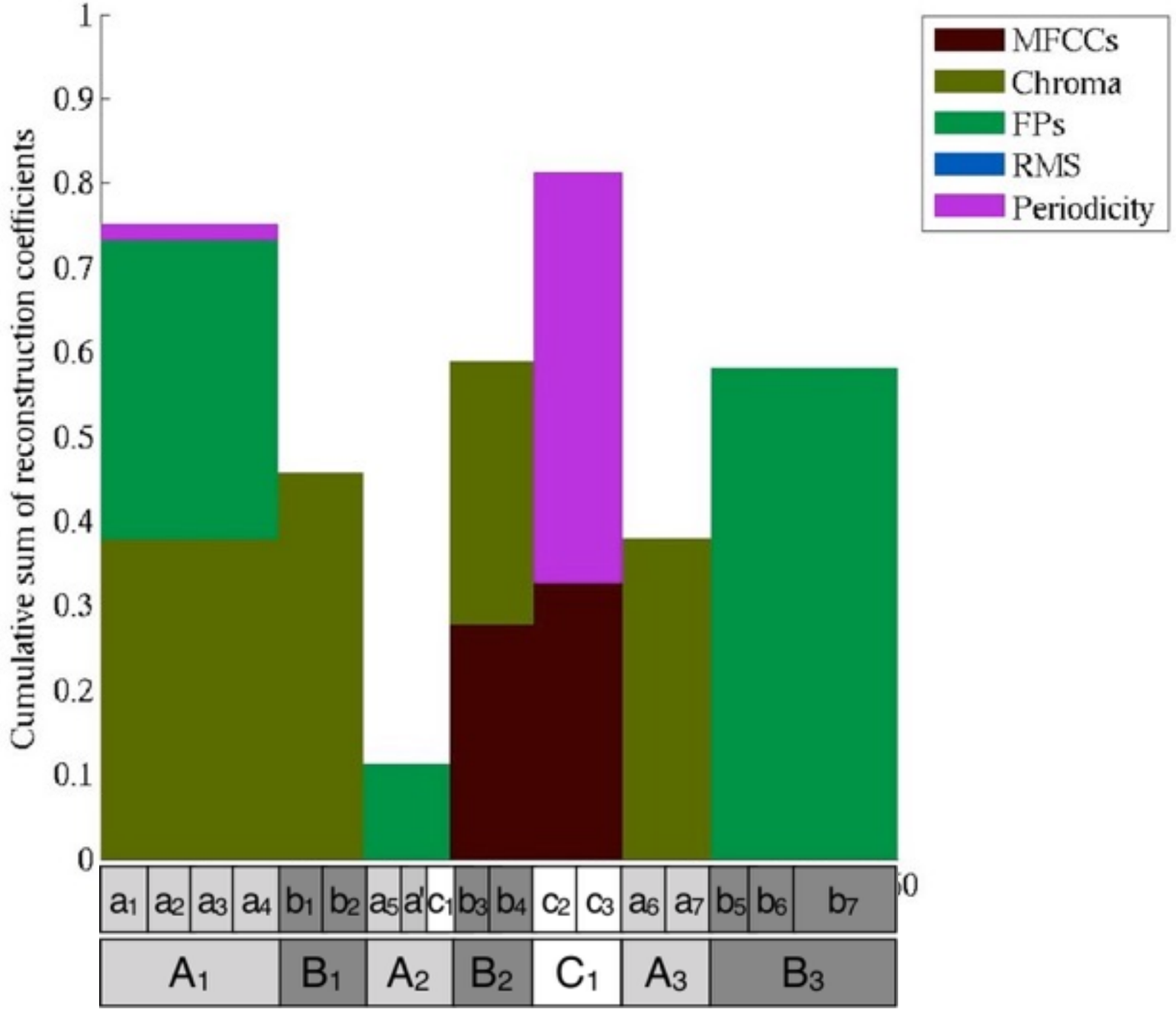
find that minimizes

Reconstruction components

$$\left(\left(\sum_{j=1}^{s}\sum_{i=1}^{f} x_{i,j}M_{i,j}\right) - N\right)^2$$

$x_{71}M_{71}$ $x_{75}M_{75}$

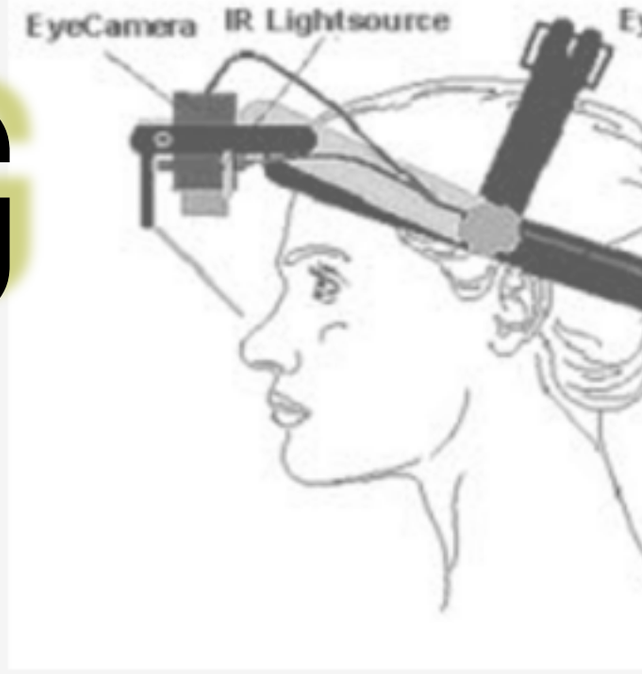
WHICH FEATURE BEST EXPLAINS THE ANALYSIS?

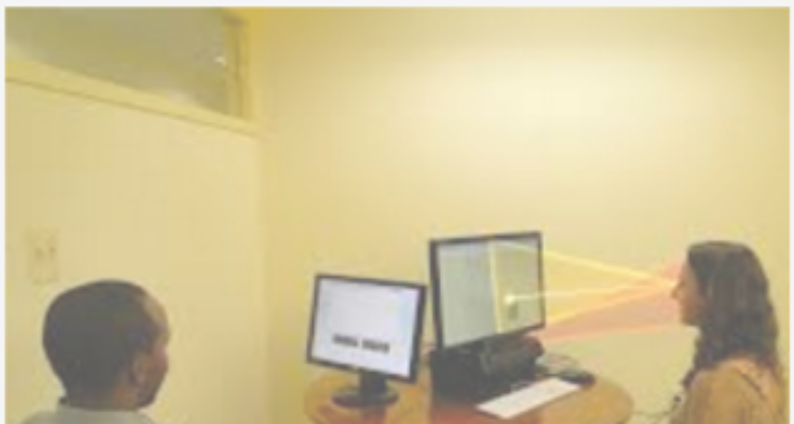


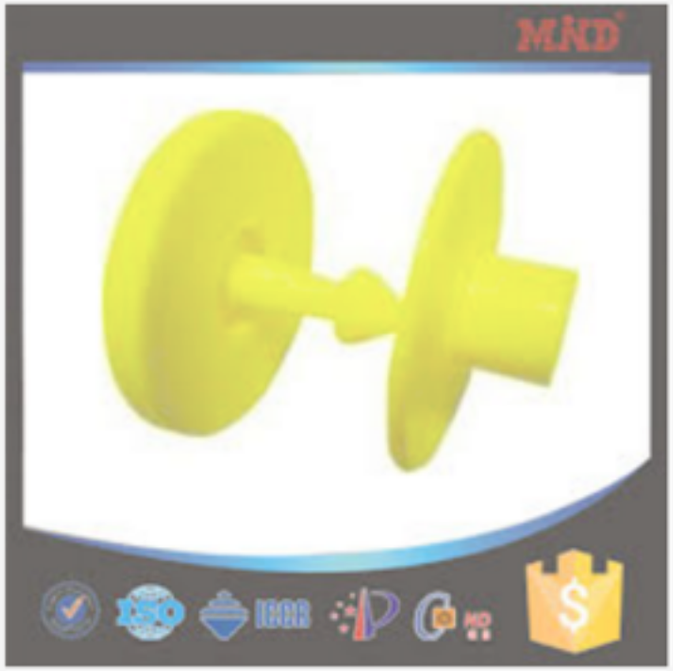
CRITICISM

- Model not validated with real-world data:
    - I.e., analyses paired with attention state of listener

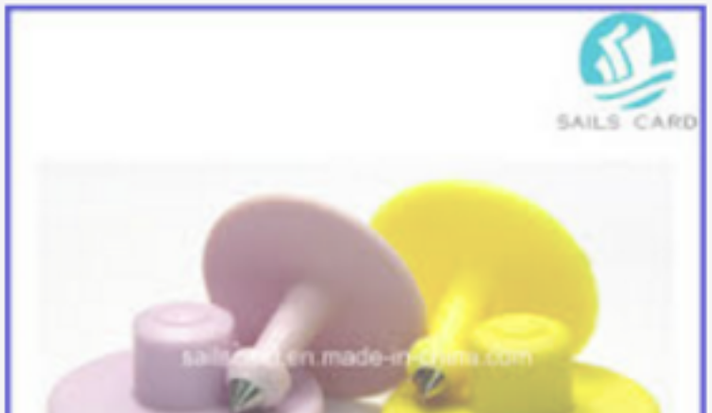
EYE TRACKING







EAR TRACKING



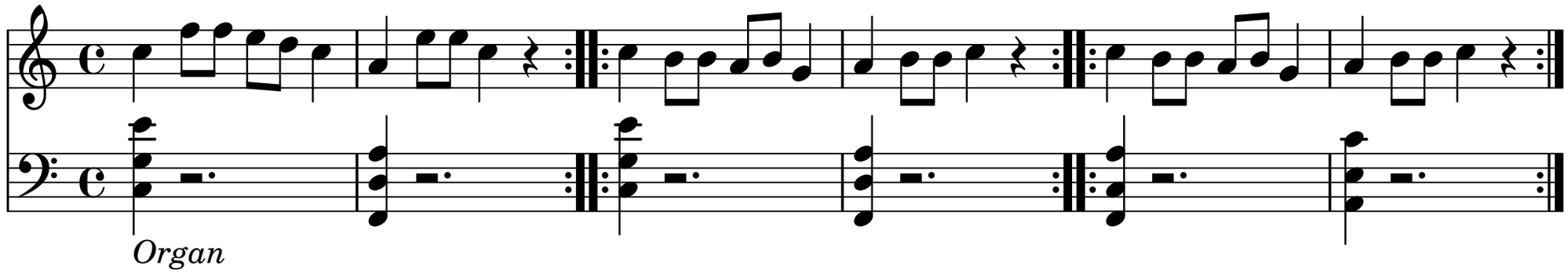
CRITICISM

- Model not validated with real-world data:
    - I.e., analyses paired with attention state of listener
    - One source of data:
        - Ran experiment testing whether focus of listener could affect perception of grouping

STIMULUS EXAMPLES

Melody ABB

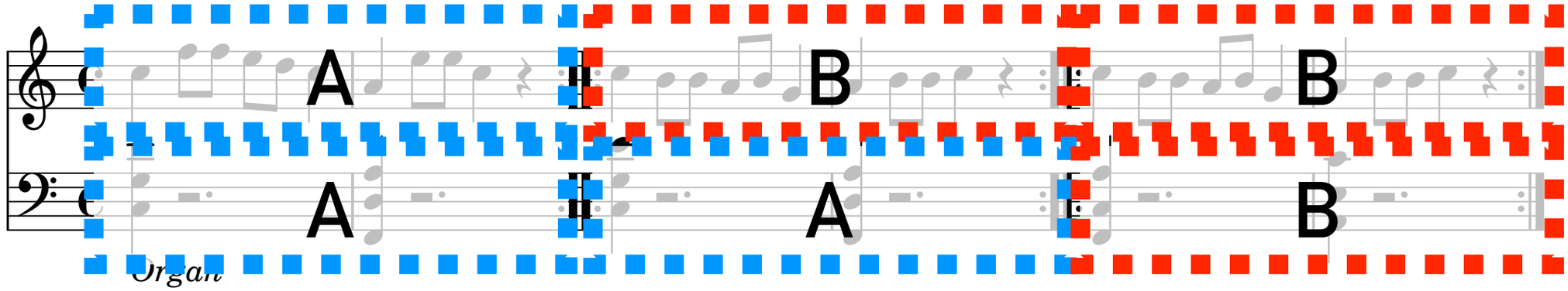
Harmony AAB



STIMULUS EXAMPLES

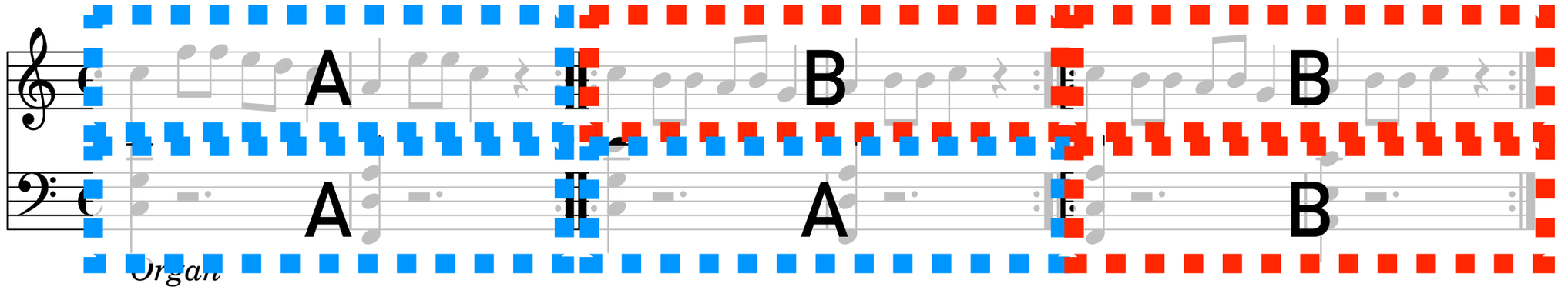
Melody ABB

Harmony AAB

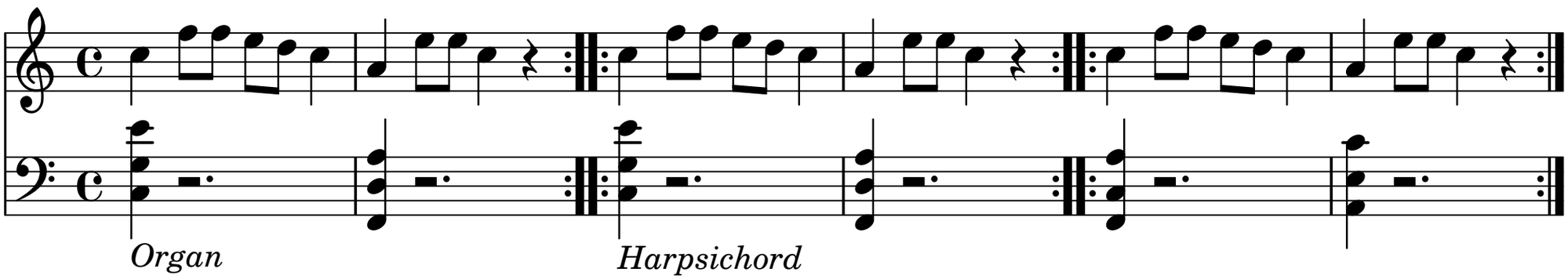


STIMULUS EXAMPLES

Melody ABB
Harmony AAB

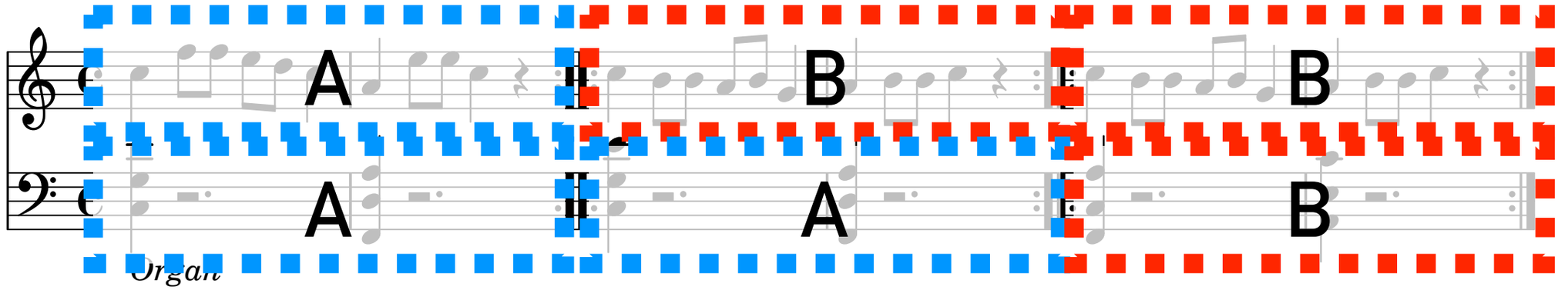


Timbre ABB
Harmony AAB

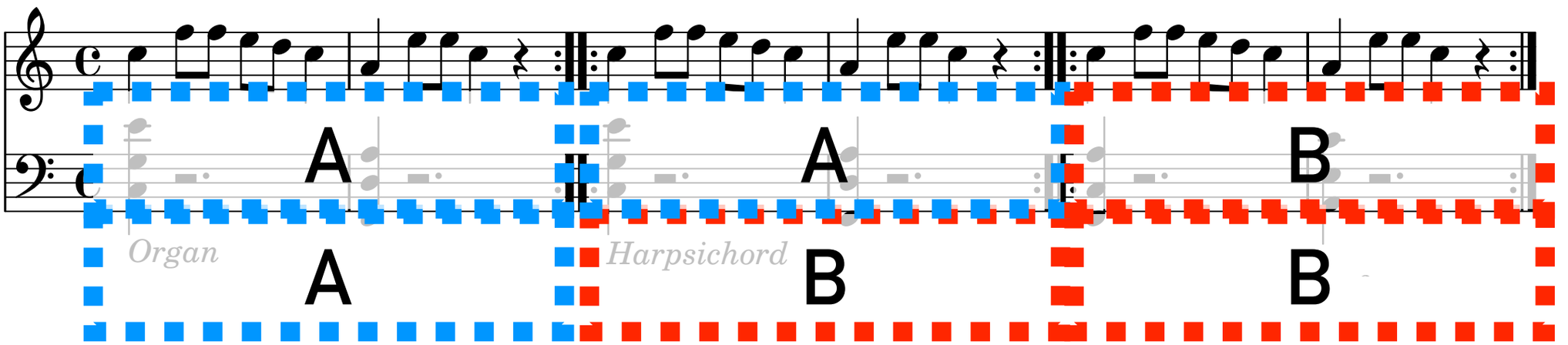


STIMULUS EXAMPLES

Melody ABB
Harmony AAB

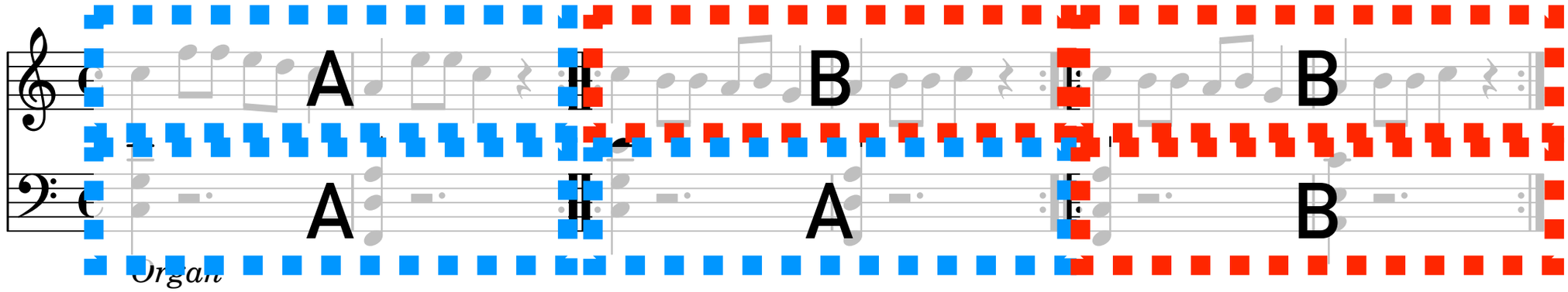


Timbre ABB
Harmony AAB

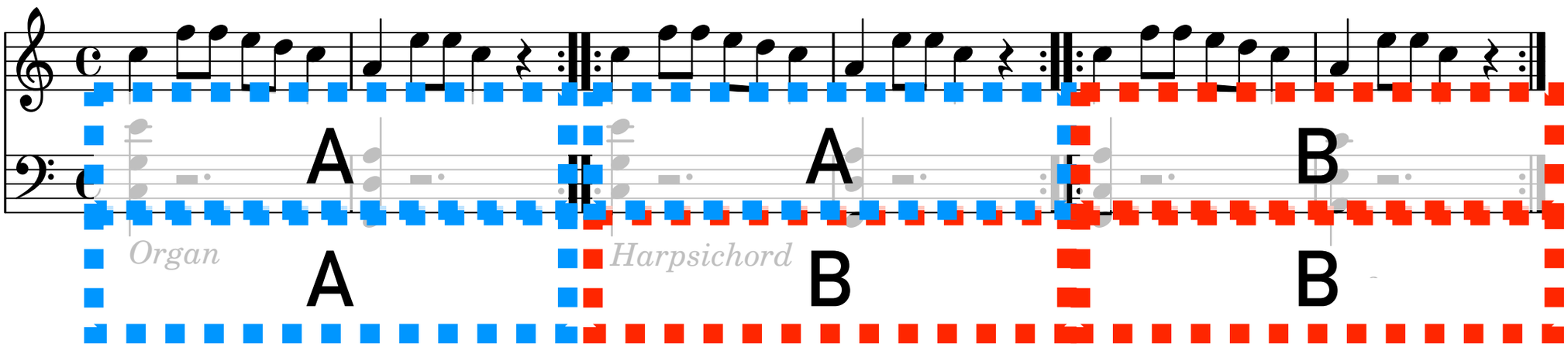


STIMULUS EXAMPLES

Melody ABB
Harmony AAB



Timbre ABB
Harmony AAB



- Experiment outcome: giving listeners a distractor task (to fix attention) affected AAB vs. ABB decisions
  - "Does this chord progression occur in the stimulus?" → prefer AAB
- Note: sometimes, organizing features are convolved

CRITICISM

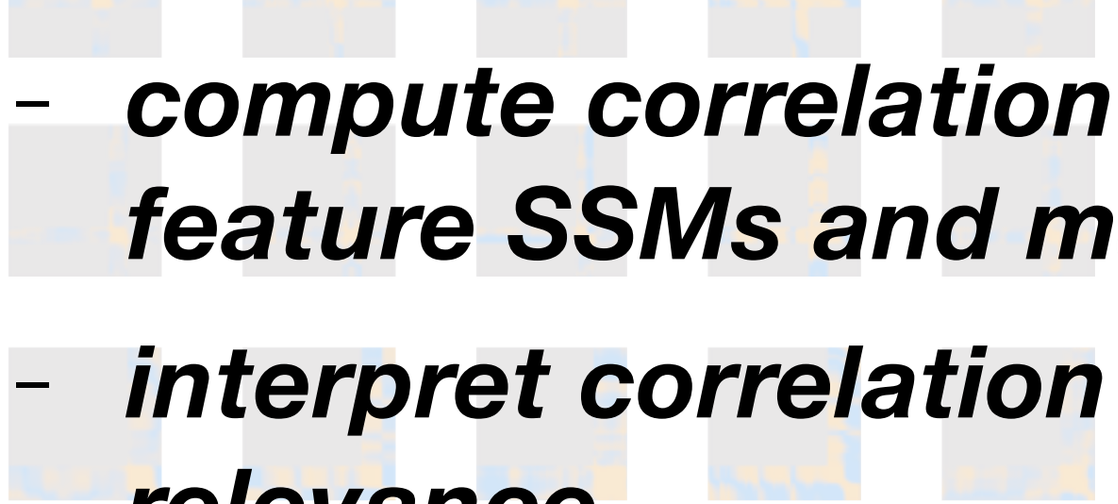
- Model not validated with real-world data:
    - I.e., analyses paired with attention state of listener
    - One source of data:
        - Ran experiment testing whether focus of listener could affect perception of grouping
- Is model needlessly complicated?
    - Try simple correlation instead

Target SSM (original annotation)

Feature-derived SSMs

Segment-based masks

Reconstruction components



QP approach:

- ***find sum of reconstruction components to generate target SSM***
- ***interpret coefficients as relevance***

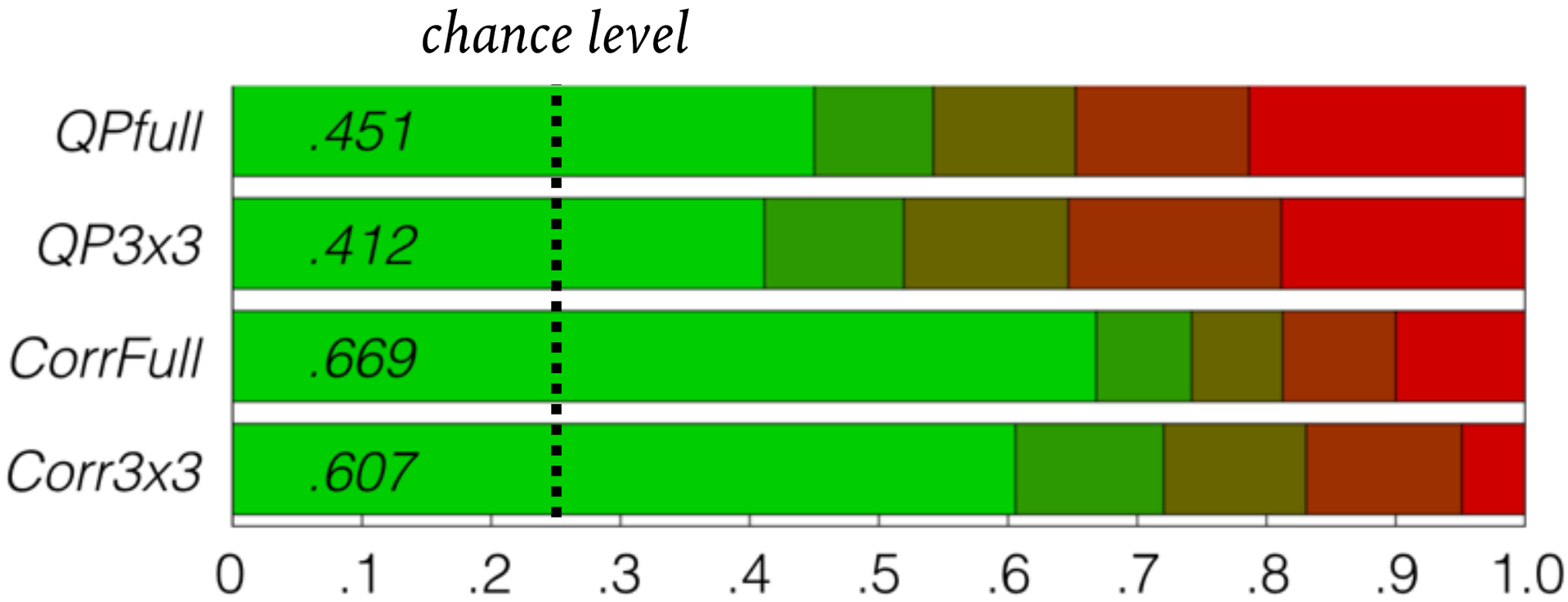
Correlation approach:

- ***compute correlation between feature SSMs and masks***
- ***interpret correlation as relevance***

EXPERIMENT

- Use algorithm to estimate relevance of features to analysis of experiment stimuli
- If the most relevant feature is the feature that changed, count correct

RESULTS



What feature was found to be most relevant?

(i.e., greatest QP weights or correlation)

SUMMARY

- Past work:
  - 1. Algorithm to estimate feature relevance from analyses (Smith and Chew, 2013)
  - 2. Artificial stimuli from psych experiment studying attention
- Current work:
  - Use stimuli (2) to validate algorithm (1) and new variations

SUMMARY

- Algorithm performs above baseline…
- …but artificial context makes non-perfect performance a disappointment!
    - Random effect of small set of stimuli
    - Mismatch between features and musicological interpretations
- Future work:
    - re-optimize (or train) model with more features, more artificial stimuli
    - apply algorithm to investigate trends in large datasets

THANKS!

ADVERTORIAL SUPPLEMENT

Imperial College Press
www.icpress.co.uk
World Scientific
www.worldscientific.com
10046 hc ISSN: 1793–0758
ISBN 978-981-3140-09-7
9 789813 140097

MATHEMUSICAL CONVERSATIONS

Mathematics and Computation in Music Performance and Composition

REFERENCES

- Jordan B. L. Smith and Elaine Chew. Using Quadratic Programming to estimate feature relevance in structural analyses of music. In *Proceedings of the ACM International Conference on Multimedia*, 113–122, Barcelona, Spain, 2013.

- Jordan B. L. Smith. Explaining listener differences in the perception of musical structure. Ph.D. thesis. Queen Mary University of London. See Chapter 6: "The effect of attention on grouping decisions."