

Design and creation of a large-scale database of structural annotations

Jordan B. L. Smith
University of Southern California
jordans@usc.edu

John Ashley Burgoyne
McGill University
ashley@music.mcgill.ca

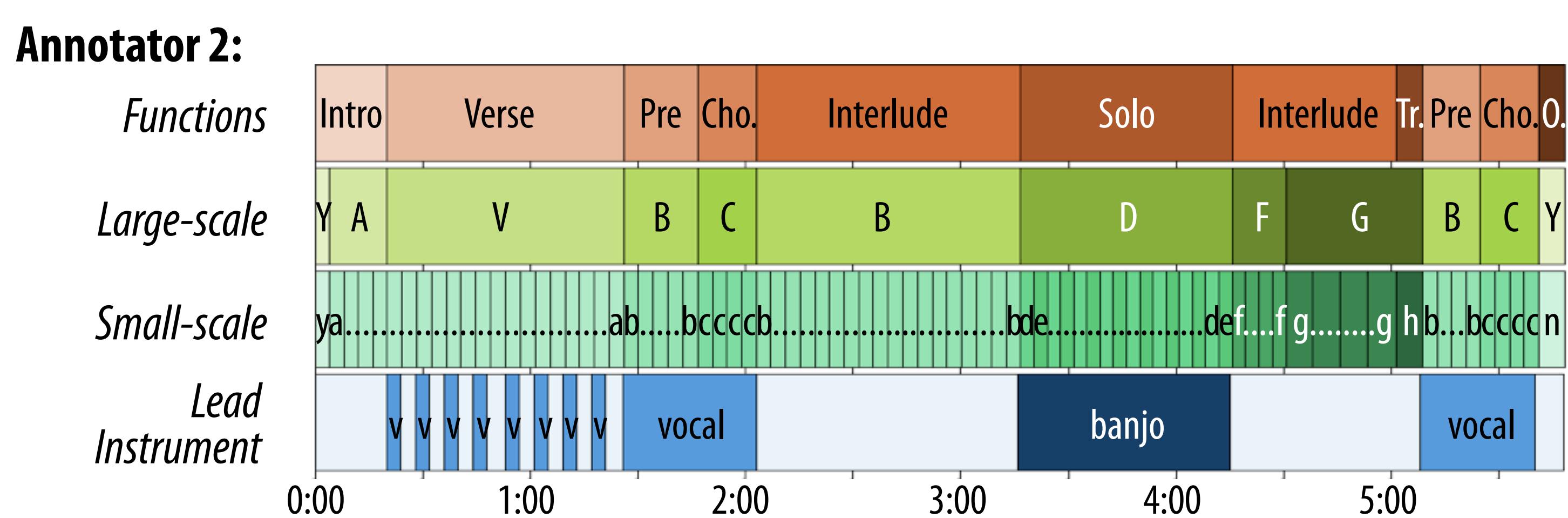
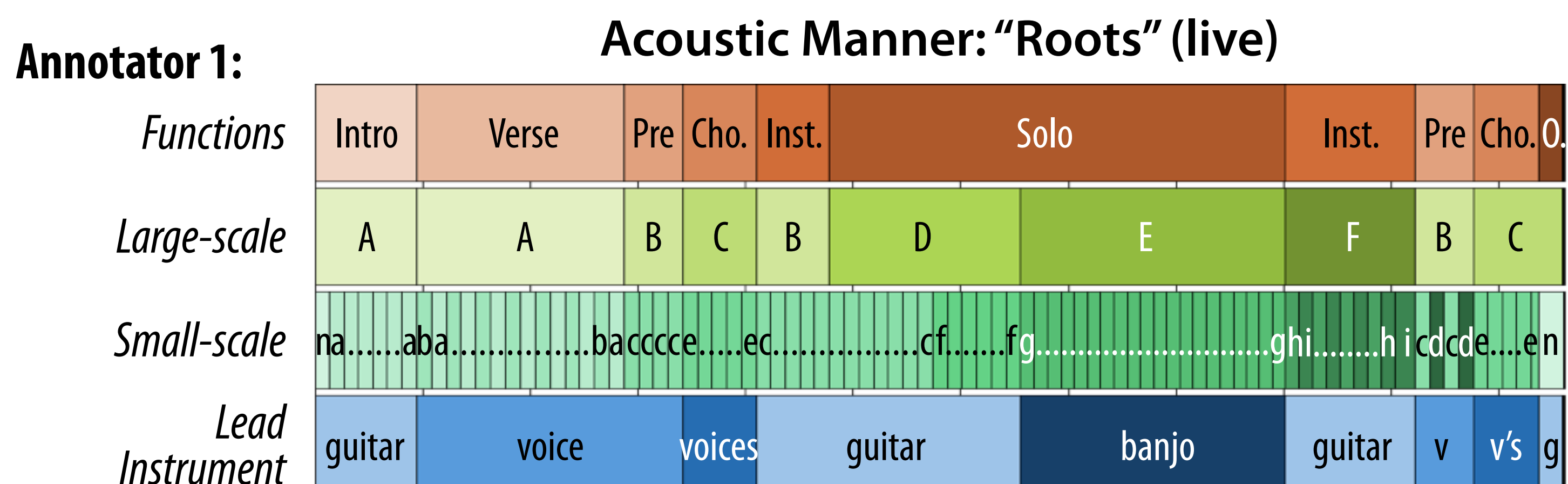
Ichiro Fujinaga
McGill University
ich@music.mcgill.ca

David De Roure
University of Oxford
david.deroure@oerc.ox.ac.uk

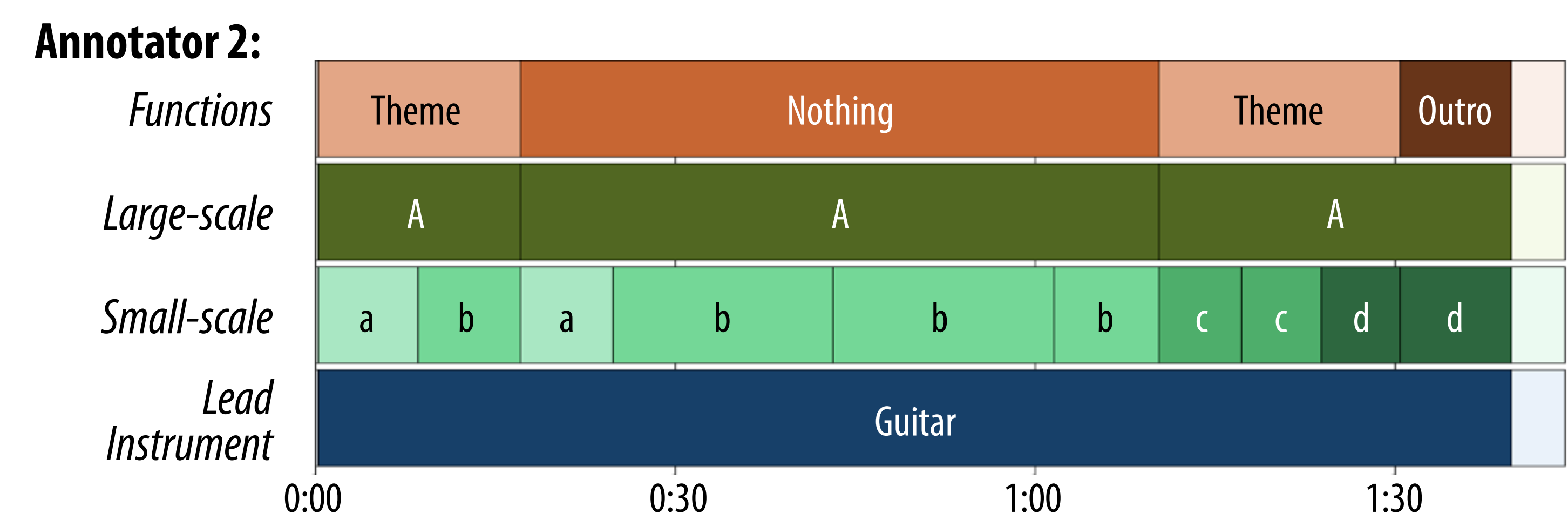
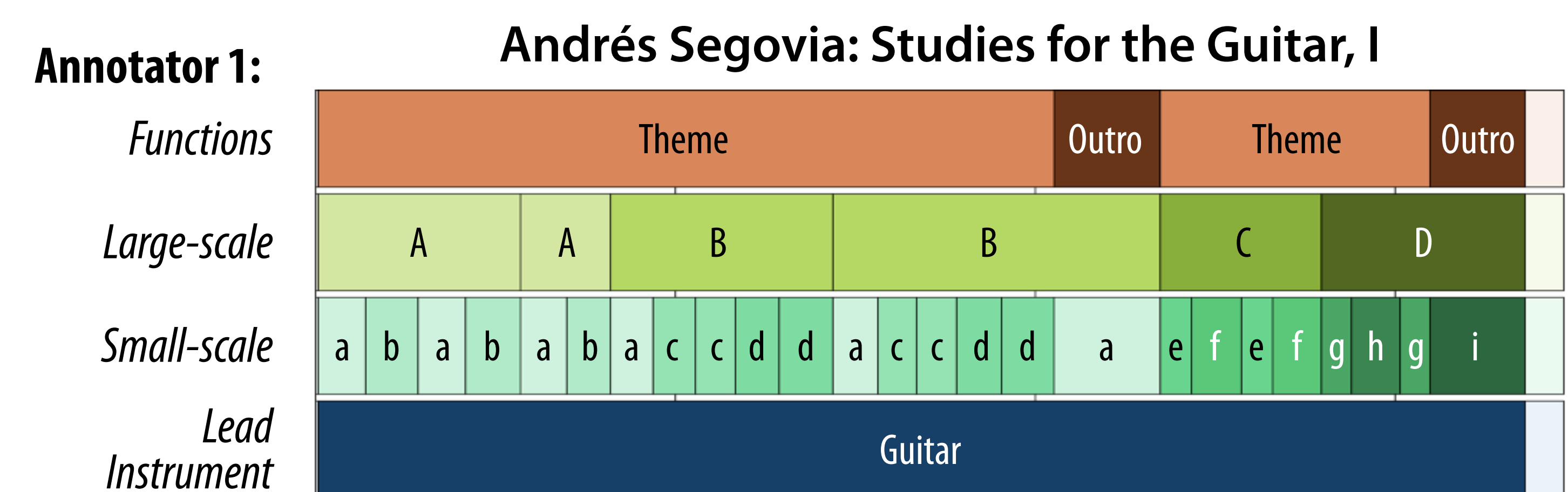
J. Stephen Downie
University of Illinois at Urbana-Champaign
jdownie@illinois.edu



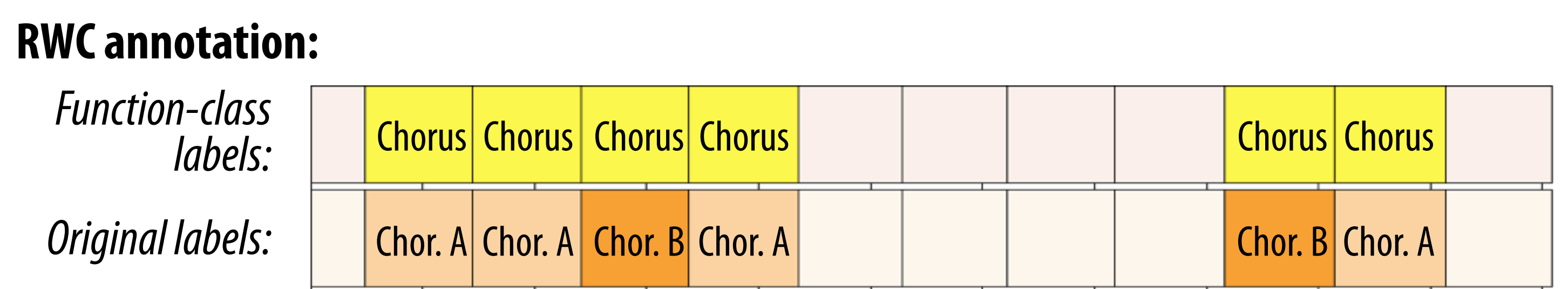
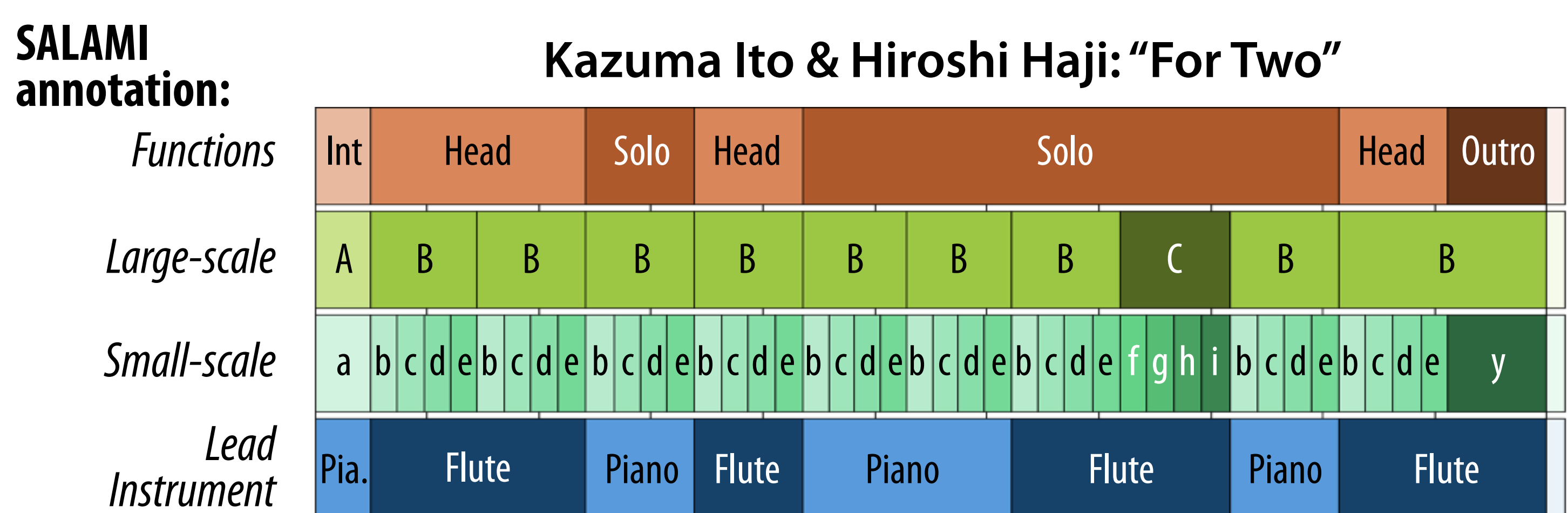
Introducing the largest collection of music structure annotations ever.



Above, a pair of annotations with average agreement between labels and above average agreement between boundaries.



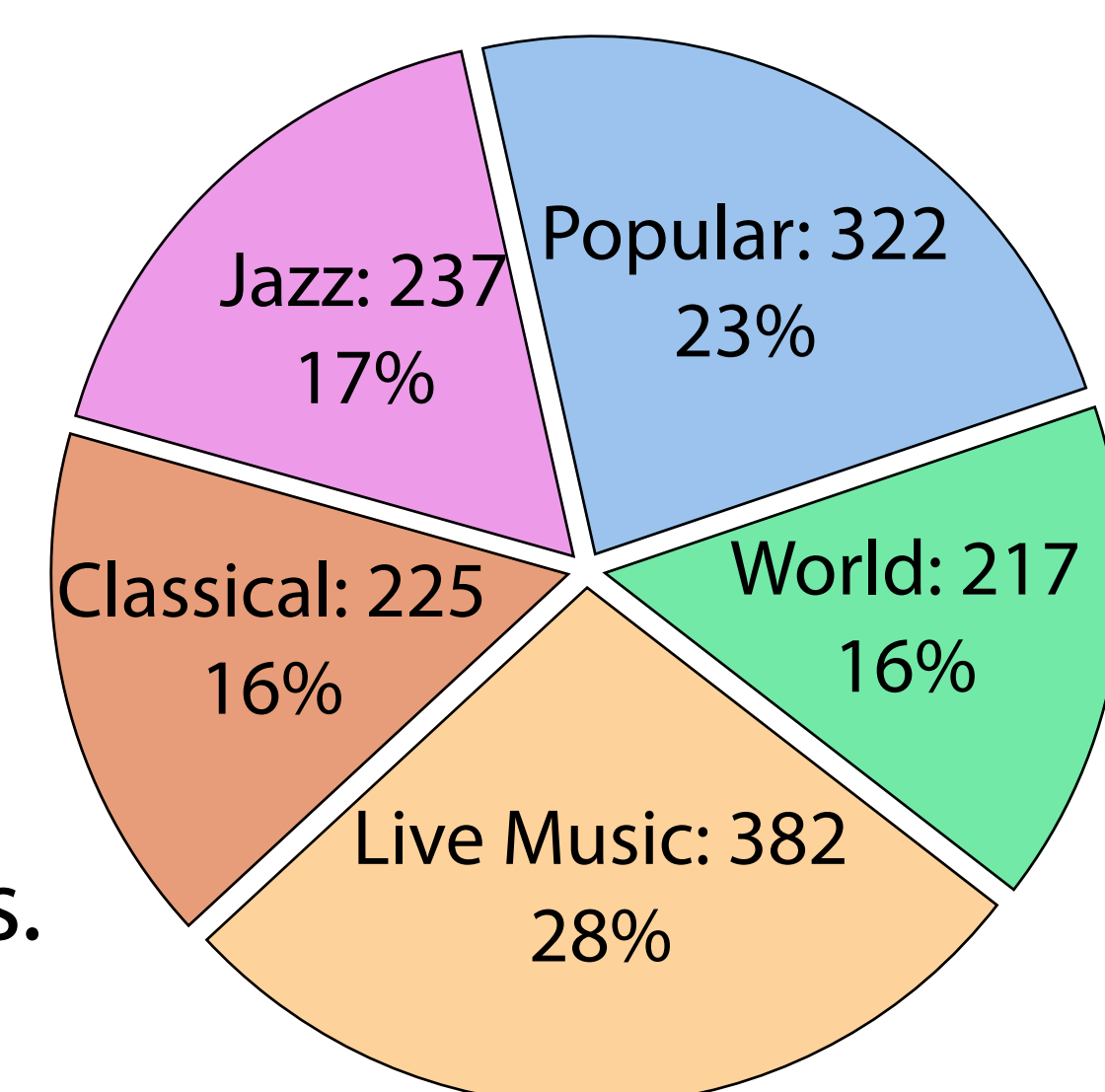
Above, a pair of annotations with poor agreement, but where the large-scale labels of the first and the small-scale labels of the second show high agreement.



Above, a single-keyed annotation compared to an RWC annotation. On average, the best match between the datasets was to take the large-scale SALAMI labels and the function-class RWC annotations.

1. The structural analysis dataset. Reinvented.

- Over 1000 double-keyed and 330 single-keyed annotations.
- Jazz, classical, popular, world music. We've got it all.
- Double-keyed annotations. Because no two people hear music the same way.
- It lets you do everything. Evaluate or train algorithms. Perform a large-scale musicological study.



2. An annotation format that is precise and rich in detail. Finally.

Musical similarity. Function. Instrumentation. All in separate layers. Inspired by Peeters and Deruty, "Is Music Structure Annotation Multi-Dimensional?" LSAS, 2009.

- Music similarity labels that are letters. Just like in music theory.
- Two levels of hierarchy. Because structure isn't one-dimensional.
- A limited list of function labels. Everything you need. Nothing you don't.
- Natural language instrument labels. Naturally.

Average # of segments per annotation:
11.2 large, 38.4 small.

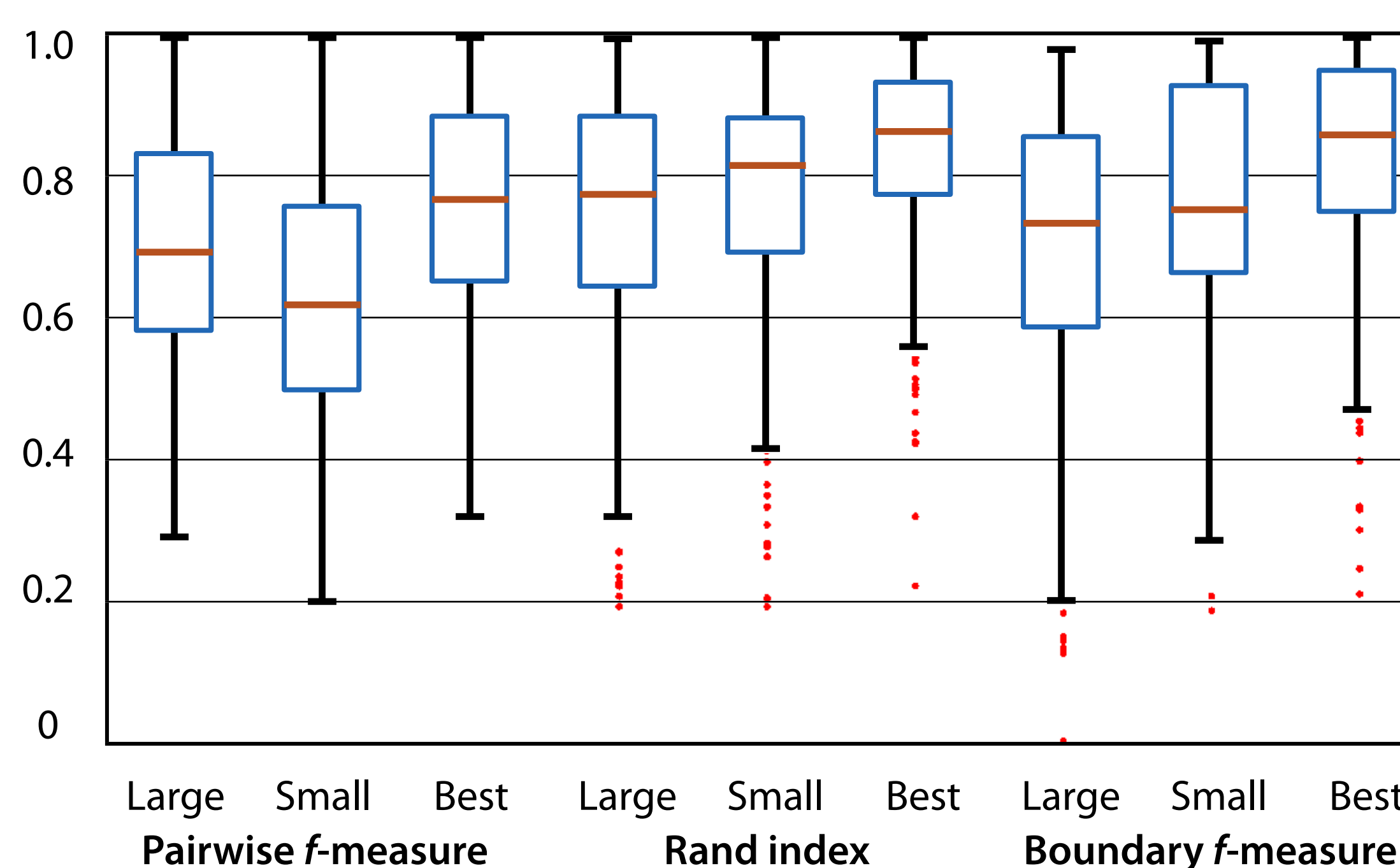
Average # of unique segments per annotation:
4.0 large, 7.2 small, 4.8 functions.

Word cloud of function label usage frequency:



3. A procedure that works like magic.

- With just a little training, annotators completed most annotations in 20 minutes apiece. Boom.
- All annotators were pursuing graduate degrees in Music Theory, Composition, or Performance.
- With two annotations and two levels of hierarchy per piece, there are four ways to compare annotations. Go ahead. You can take the best match.



Box and whisker plot of average inter-annotator agreement between pairs of large-scale analyses, pairs of small-scale analyses, and the best alignment found between any pair (including one annotator's large and the other's small).

Everyone should have access to such a great database. And now everyone can. Make the switch to **SALAMI**. Learn more about the **Structural Analysis of Large Amounts of Musical Information** project at salami.music.mcgill.ca.

