

SALAMI: The Structural Analysis of Large Amounts of Music Information project

SALAMI in a nutshell:

Our goal is to provide an unprecedented number of structural analyses of pieces of music for future study. These studies may include training computer programs to automatically do structural analysis themselves, or tracing the evolution of form over centuries, or investigating which forms seem to dominate which genres, or things of that nature. In light of this, we are striving to cover a wide variety of musics, from western popular to Indian classical, including recordings in live and studio settings.

Structural Annotators:

Your job will be to generate the structural descriptions of pieces of music that will be used in this research, and you will strive for both accuracy and speed. Of course, analyzing the structure of a piece of music is hard: it requires skill and judgment, and it can't usually be said that there's a "right answer" for a particular piece. It's also often a very fuzzy process: the exact definition of "form" can be hard to pin down, and even musical processes that are relatively well defined (e.g., a modulation) can be very tricky to locate in the music. [Along these lines, please put quotation marks around any word in this pamphlet that you think is being abused.] Despite this, the analyses you produce will need to be very strictly laid out (in fact, they will need to be expressed in a machine-readable way), in a way that is consistent across annotators. This need for consistency is part of the reason we are holding this training/audition today.

Definition:

What do we mean by "formal analysis"? To start with, since many of you are music theorists, here are some examples of what we don't mean:

- A classification of the piece into a formal type such as sonata, song, or canon.
- A Schenkerian reduction of the piece into its *Ursatz*.

What we do mean can be roughly expressed as: "the organization and division of [the piece] into definite sections, and the relation of those sections to each other." (This is Salzer's definition of 'form' as distinct from 'structure' and 'design,' excerpted from the Oxford Dictionary of Music.) Put another way, we will have you partition each piece into several segments, and then give these segments appropriate labels to describe which are similar to each other or which fulfill a related musical role. This definitely overlaps with the first definition above, the "classification of the piece," except that for a given rondo we wouldn't want you to produce the answer "rondo" but the answer "ABACABA."

How to analyze music:

Before the procedure you will use for annotating music is outlined, it may help to consider a few different approaches one could take to do this kind of segmentation + labelling analysis.

1. One approach could be for you to first determine what the formal type of the piece you are listening to is (e.g., sonata, rondo, pop song), and then segment and label the piece according to your knowledge of that formal type (e.g., using the definitions of ‘exposition,’ ‘development,’ and ‘recapitulation’). However, lots of the music you listen to will be unfamiliar to you and you may not know what formal type (and corresponding vocabulary) to apply; besides which, many pieces of music simply cannot be explained with a single formal type.

2. Another approach could be purely perceptual: you would listen for prominent harmonic or rhythmic boundaries in the piece in order to segment it, and apply labels to the resulting sections by comparing them and determining which were similar to each other. One problem with this approach is that it does not reflect the function of different segments (e.g., an ‘introduction’ section and a ‘transition’ section might have exactly the same music, but embody different structural roles). Additionally, it is impossible to define precisely what ‘similar’ means. [Note: this will still be a problem in the system proposed later, but it bears mention now.]

3. An alternative to all this could be to simply report what occurs in the music itself, by making note of the instrumentation, of major melodic ideas, of portions of relative silence or rhythmic confusion, and so forth. But this effort at objectivity may quickly lead you away from structural analysis and towards a full transcription, which we do not want.

The annotation method we have in mind (described in the next section) draws on aspects of all these three types of analysis listed above, but has the advantage of not conflating them. By separating the organization of instrumentation, musical material, and formal function, the method allows you to analyze a huge variety of pieces of music using a single, highly-constrained vocabulary.

Finally, it should be noted that none of the three methods mentioned above seem to do anything to address the fact that musical structure is frequently hierarchical. When analyzing a piece, it can be hard to know what timescale is appropriate: a single description of a piece such as “ABABCAB” might not reflect that each A is composed of two contrasting parts (= “ADBADBCADB”), or that the sequence “AB” has significance at a larger scale (= “DDCD”). Again, although this is an ambiguity that no system can completely resolve, the method described below includes a few markers that will partly address this.

SALAMI Annotation Procedure:

As stated before, the structural analysis we want of each piece will consist of a partitioning of that piece into sections, and the labelling of these sections. However, this partitioning will happen on three independent levels:

1. The level of acoustic similarity;
2. The level of musical function;
3. The level of instrumentation.

For each level of structure, the labels available to you are:

1. Acoustic similarity:

A, B, C, D, E, ... : these indicate musical phrases, ideas, or subjects that may be differentiated on the basis of rhythmic, melodic, or harmonic material. The idea is that each particular musical idea gets its own label. We advise limiting your annotation to 5 labels, but if you truly require more letters you are permitted to use *F, G, H*, and so on. (Although you will see below, two markers: *X* and *=*, may obviate the need for more than 5 letters.) Note that every instant in the piece must be labelled with a letter.

Z is a special letter that is used to denote a complex or potentially amusical section that stands out strongly from the rest of the piece. For instance, a sudden free improvisational part, or a brief spoken dialogue in the middle of a piece might be appropriately labelled with *Z*. Note that in cases where a piece has two such inscrutable sections, they should both be labelled as *Z* even if they are not acoustically similar. *Z* is an exception among letter labels in this respect.

' : the prime symbol (marked for our purposes with an apostrophe) is commonly used in describing the structure of classical works, to indicate when a particular section occupies a gray zone between being a repetition of a previous musical idea, and being a new, independent musical idea. It could be called for, for instance, if a particular passage were repeated, retaining its musical identity but being transposed and converted from the major to the minor mode.

2. Musical function:

Depending on the piece under analysis, many words could potentially be used to describe the function of a particular segment. The vocabulary used will depend on the genre of the piece being analyzed. However, in the interest of generating consistent analyses, we strongly encourage you to use terms from the following list where they seem to be reasonably appropriate. Note that not every moment in a piece needs to have a function label.

- *intro*: a part that leads in to the rest of the piece.
- *outro*: a part that leads out of the rest of the piece. Also known as a “coda.”
- *transition*: a part that helps transition from one segment to another. In various contexts, a transition may be known as a “Bridge,” “Middle eight,” or by other phrases, but “Transition” is preferred here.

- *solo*: a part in which a single instrument or voice comes to the foreground.

Genre-specific terms:

- *verse*: in a song, a section in which the tune remains the same, but the text changes with each repetition.

- *chorus* (aka refrain): in a song, a part which contrasts with the verse and which is repeated more strictly. Sometimes two distinct chorus-like sections are present in a single song; in this case, they should still both be labelled as *chorus*, since they will be distinguished by different letter labels.

- *head*: in a jazz piece, a statement of the theme on which the piece is based.

- *main theme, secondary theme*: in a classical repertoire, these may indicate two principle contrasting musical ideas.

- *exposition, development, and recapitulation*: a specialized vocabulary for sonata form.

- *variation*: this label may be required if the main organization of a piece is a theme followed by a series of variations.

In the course of annotating thousands of pieces, we expect that the above list may prove incomplete, so although we present the above lists as explicit suggestions, you are encouraged to let us know if you believe a new term should be introduced. Then we can communicate that to all the annotators to keep people on the same page.

3. Instrumentation:

In polyphonic contexts, those segments where there exists a main melodic referent should be labelled with the appropriate instrument. Here the vocabulary is in principle unrestricted: simply label the segment with the instrument or voice that contains the melody.

For instance, in a rock song, you would label those segments featuring the lead vocal with *vocal*. In other parts of the same song, if the guitar takes the melody, label those sections *guitar*. Where two distinct voices use the same instrument, append a number to the label. Sections with different instrument labels may potentially overlap: for instance, in a duet, *vocal1* and *vocal2* may sing certain portions together and others separately.

4. Special markers:

X and *=*: these are two special markers which apply to instants in time, as opposed to all the previous labels, which apply to time segments. They are used to indicate that a particular segment is subdivided into two parts, with either the two parts being identical or highly similar to each other (*=*) or highly dissimilar (*X*). They are always placed in reference to the lettered segments. Note that a single lettered segment may have multiple markers within it.

Here we summarize some features of this labelling system:

Exclusion rules:

Each segment of the music may be tagged with as many labels as apply to the segment. However, some labels cannot be simultaneously applied to the same moment in time. To summarize:

- Only one letter label (*A, B, C, D, E, ..., Z*) may be applied to any particular segment of a piece. (Recall also that every moment in a piece *must* have a letter label applied to it.)
- Only one function label may be applied to any particular section.
- Leading instrument labels may overlap as necessary.
- Only one marker (*X* or *=*) may be applied to a single instant.

Equivalence rules:

Some labels indicate that the segments to which they have been applied are ‘equal’ in some sense. For instance, all segments with the label *C* should contain the same musical idea. Additionally, all segments with a particular instrument label should actually feature the same instrument. (Where they don’t, it was specified above to use numerical markers: thus the two voices in a duet are labelled *vocal1* and *vocal2*.) On the other hand, it is not assumed that any two segments with the same function label must actually sound the same in any way; they merely have the same function. Thus, several different-sounding segments that all have transitioning function may be labelled with *Transition*, even if they do not resemble one another.

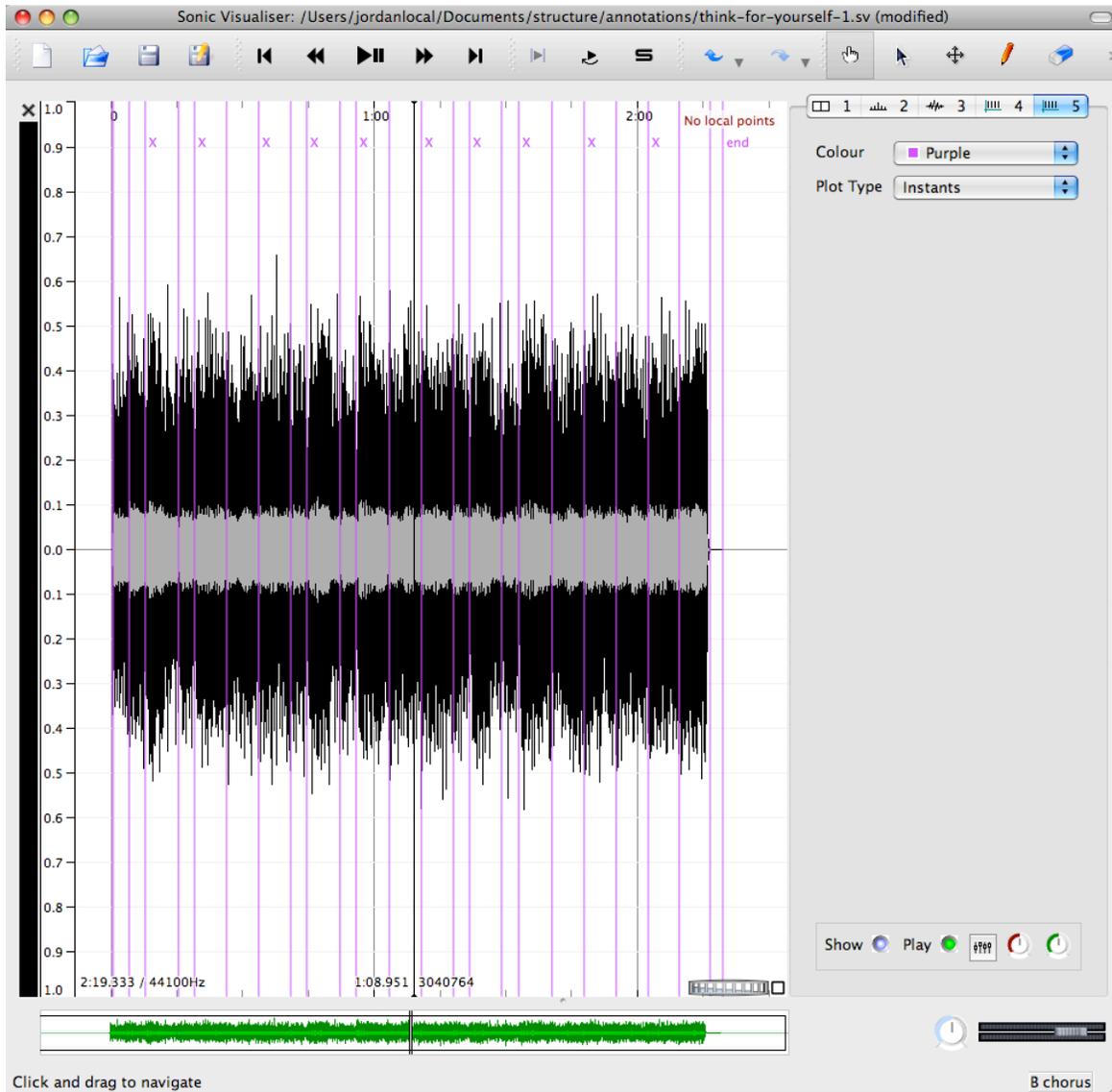
The main exception to this rule is the label *Z*, since a song may contain several uncategorizable segments that do not sound alike. The following table summarizes which segments obey an ‘equivalence’ rule and which do not:

Labels which denote equivalence:	Labels which may not denote equivalence:
All letter labels (<i>A, B, C, D, E, ...</i>) All instrument labels (<i>piano, vocal, ...</i>)	<i>Z</i> All function labels

Examples:

The purpose of the training session on Thursday, May 6th is to go over several examples of how this system may be applied to annotate particular songs, and to demonstrate the software, Sonic Visualizer, that will be used to perform this task. However, to give a flavour of how this all works, we present a brief example annotation for the song “Think For Yourself,” by the Beatles, which you can listen to here:

<http://www.youtube.com/watch?v=yXGSBgr8sbg>



Above is a screenshot of the user interface for Sonic Visualizer, zoomed out so that the entire song is in view. The annotated boundaries are indicated as vertical purple lines, and labels for some of these boundaries are given at the top.

Below, you'll find the completed structure description for this song. The time of each structural boundary is given in the first column in seconds. The label for that boundary is given in the second column. Each segment to which each label applies thus extends from the time given to the left of the label up to the time given on the next line that defines a new label. For a label to extend over multiple lines, explicit begin/ and /end tags (such as *vocal/* and */vocal*) are given for that label.

For example: the second *verse* extends from 15.15 to 26.28 seconds, with an *X* marker at 18.89 seconds.

Think For Yourself
The Beatles

0.000000000	silence
0.429569160	C intro
4.109931972	A verse vocal/ guitar/
7.783401360	X
15.153628117	A verse
18.890045351	X
26.284988662	B chorus
33.523809523	X
40.857528344	A verse
44.563242630	X
52.006893424	A verse
55.805986394	X
63.245351473	B chorus
70.530612244	X
77.929659863	A verse
81.653061224	X
89.042721088	A verse
92.834535147	X
100.327619047	B chorus
107.647709750	X
114.956190476	B chorus
122.258730158	X
129.544126984	= /vocal /guitar
136.434648526	silence
139.334648526	end