Musical Grouping Structure

David Meredith

Department of Computing,

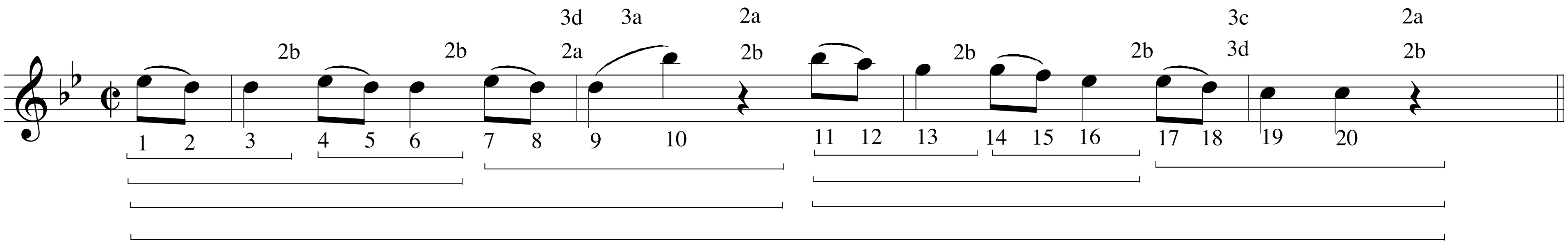
City University, London.

dave@titanmusic.com

www.titanmusic.com

MSc/Postgraduate Diploma in Music Information Technology Lecture
Department of Music, City University, London
Friday, 7 March 2003.

1. Musical grouping structure

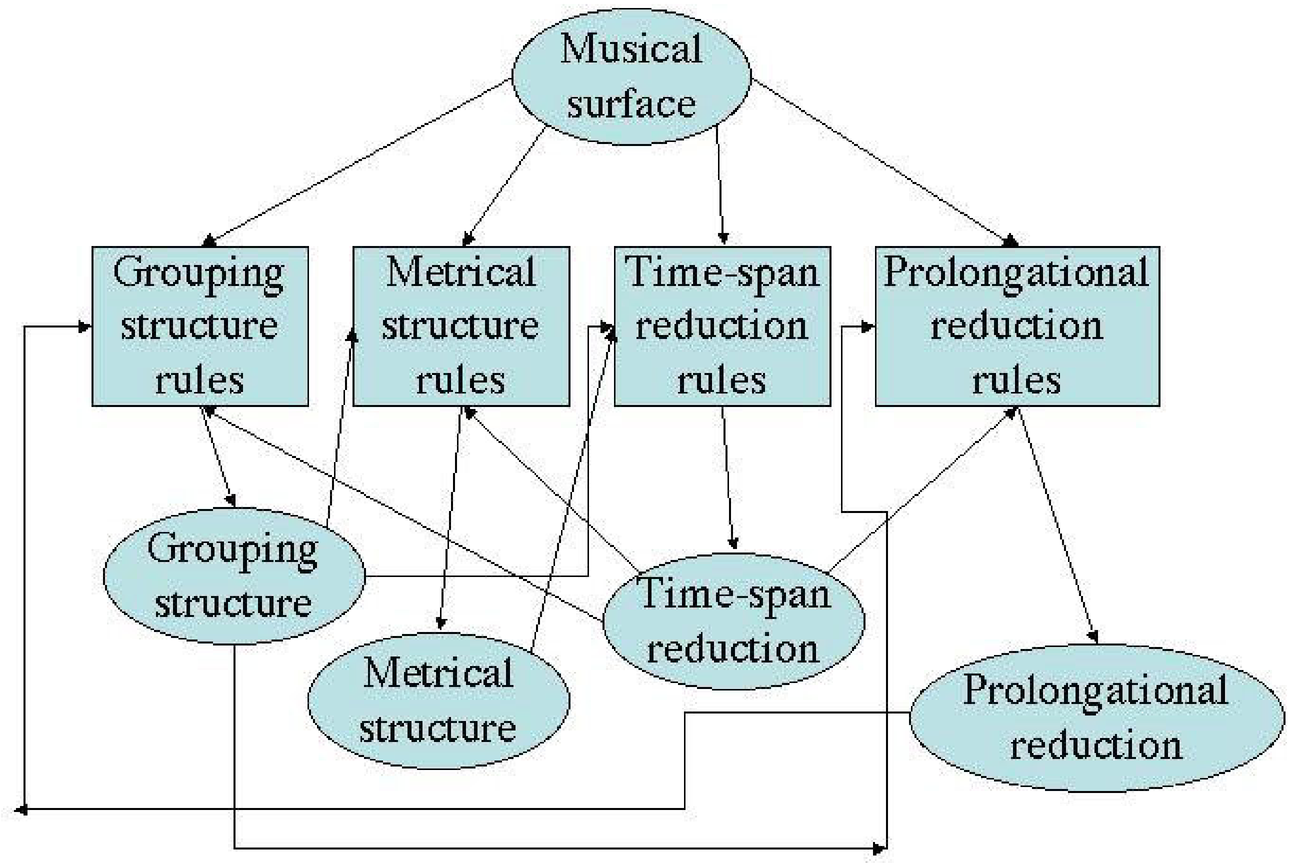


1. Musical grouping structure

1. Lerdahl and Jackendoff (1983, p. 12) point out that when we listen to a passage of music we automatically segment or chunk the sound we hear into units of widely varying sizes such as motives, themes, phrases, theme-groups, sections and so on.
2. Lerdahl and Jackendoff (1983) use the term *group* to refer to such structural units.
3. Certain types of groups are typically explicitly notated in scores by means of phrase marks, slurs and breath marks—singers and players of wind instruments typically breathe *between* groups rather than in the middle of them.
4. Also, a number of experiments have shown that performers tend to slow down at the boundaries between certain types of groups such as phrases and major sections in order to better convey the structure of the music.
5. The *grouping structure* of a passage of music is therefore the way that the piece is perceived to be cut up or chunked into segments that form structural units of various sizes.
6. Lerdahl and Jackendoff (1983, p. 13) go so far as to claim that "grouping can be viewed as the most basic component of musical understanding" because once a listener has "construed a grouping structure" for a passage, he or she "has gone a long way towards 'making sense' of [it]".
7. To give you a better sense of what is meant by grouping structure, consider this melody here [Mozart G minor symphony] which comes from the beginning of Mozart's G minor symphony. This fragment sounds like this [PLAY Mozart G minor symphony].
8. The first thing to notice is that this particular fragment sounds reasonably complete in itself. In other words, the whole fragment forms a structural unit or group.
9. On the next lower level, most of you will probably agree that the melody can naturally be perceived as being constructed from two phrase-type groups: an antecedent phrase that ends here, followed by an answering phrase that ends here.
10. Many of you might even perceive each of these two phrase-level groups to be constructed from yet smaller groups that correspond to what music analysts typically call motives. For example, you might perceive this opening phrase to be composed from three structural units as marked here.
11. In this lecture I'm going to introduce you to some of the more important theoretical, experimental and computational work that's been done on *musical grouping structure*.
12. Like the theories and models of metrical structure that I talked about last week, each of the theories of grouping structure that I'm going to describe this week takes the form of a system of rules, an algorithm or a computer program that accepts a representation of a musical passage as input and generates as output a structural description that is supposed to correctly describe the grouping structure that is heard by an experienced listener when he or she listens to the passage.

13. For example, if you gave such a theory as input a digital audio recording or a MIDI file representing a performance of this melody here, we would want the model to generate as output a representation of the grouping structure of the melody that we've indicated here in this diagram using phrase marks or slurs.

2. Lerdahl and Jackendoff's (1983) *Generative Theory of Tonal Music*



- WELL-FORMEDNESS RULES define CLASS of possible structural descriptions
- PREFERENCE RULES used to find BEST structural description

2. Lerdahl and Jackendoff's (1983) *Generative Theory of Tonal Music*

1. I'm going to start with Lerdahl and Jackendoff's (1983) theory of grouping structure which forms one component of their *Generative Theory of Tonal Music* or GTTM.

2. As you may recall from last week, *GTTM* is essentially a system that takes a representation of a passage of music as input and generates as output four structural descriptions of the passage that are supposed to correctly describe certain aspects of how the passage is interpreted by an expert listener.

3. *GTTM* consists of four interacting components:

   (a) the first of these generates a grouping structure;

   (b) the second component deals with metrical structure—I discussed that part of the theory in detail last week;

   (c) the third component generates what's called a 'time-span reduction' which is supposed to describe the way that certain events in the passage are heard as being elaborations of other events; and

   (d) the fourth component of the theory uses the time-span reduction of a passage to generate a 'prolongational reduction' of the passage which is intended to represent the perceived 'ebb and flow' of tension and relaxation created by the harmonic structure of the passage.

4. Each of these four components consists of two sets of rules:

   (a) a set of *well-formedness rules* that generatively define a class of structural descriptions; and

   (b) a set of *preference rules* that are designed to work together to isolate the structural description that best describes how the listener interprets the passage given to the theory as input.

3. Lerdahl and Jackendoff's (1983) theory of grouping structure

- Musical grouping is an "auditory analog of the partitioning of the visual field into objects, parts of objects, and parts of parts of objects" (Lerdahl and Jackendoff, 1983, p. 36).
- Theory of grouping "seems to consist largely of general conditions for auditory pattern perception that have far broader application than for music alone" (Lerdahl and Jackendoff, 1983, p. 36).
- "A listener needs to know relatively little about a musical idiom in order to assign grouping structure to pieces in that idiom" (Lerdahl and Jackendoff, 1983, p. 36).
- Theory can only cope with homorhythmic music.

3. Lerdahl and Jackendoff's (1983) theory of grouping structure

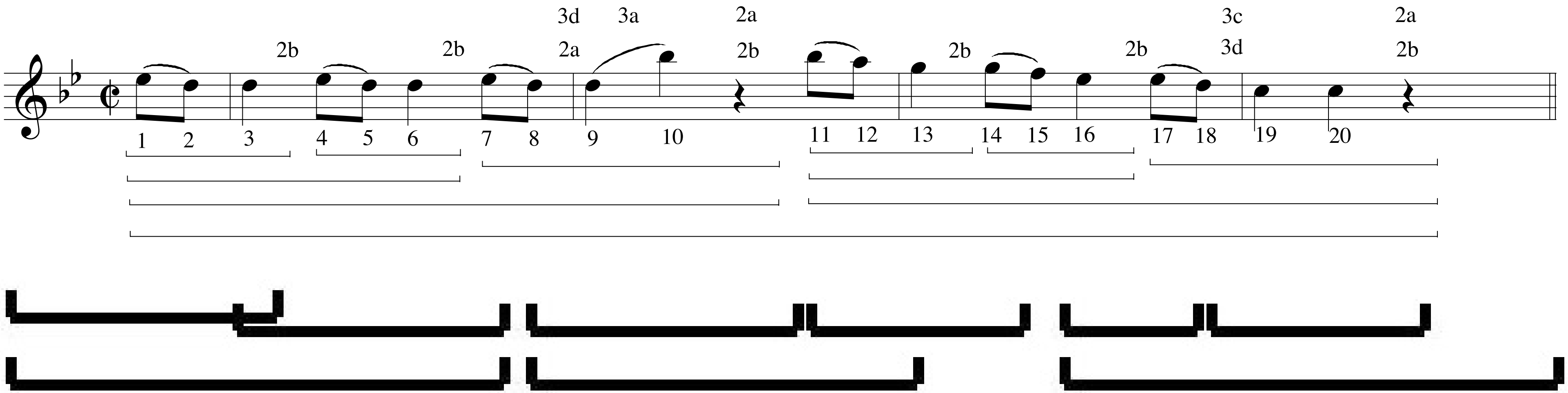
1. Lerdahl and Jackendoff (1983, p. 36) suggest that "grouping of a musical surface is an auditory analog of the partitioning of the visual field into objects, parts of objects, and parts of parts of objects." and they go to some lengths to demonstrate that the rules the seem to govern the way that we chunk up music into groups are very much like those that seem to govern the way that we parse visual scenes into objects.

2. They also point out that, unlike the other components of their theory, Lerdahl and Jackendoff's (1983) theory of grouping "seems to consist largely of general conditions for auditory pattern perception that have far broader application than for music alone."

3. In particular, some of their grouping rules are idiom-independent—that is, they can be used to predict the grouping structures of passages in musical styles other than Western classical music. Or, to put it another way, "a listener needs to know relatively little about a musical idiom in order to assign grouping structure to pieces in that idiom" (Lerdahl and Jackendoff, 1983, p. 36).

4. An important caveat to remember, however, is that their theory only works on homorhythmic music—that is, music in which every part has the same rhythm.

4. Lerdahl and Jackendoff's (1983) Grouping Well-Formedness Rules



GWFR 1 Any contiguous sequence of pitch-events, drum beats, or the like can constitute a group, and only contiguous sequences can constitute a group.

GWFR 2 A piece constitutes a group.

GWFR 3 A group may contain smaller groups.

GWFR 4 If a group G_1 contains part of a group G_2, then it must contain all of G_2.

GWFR 5 If a group G_1 contains a smaller group G_2 then G_1 must be exhaustively partitioned into smaller groups.

4. Lerdahl and Jackendoff's (1983) Grouping Well-Formedness Rules

1. Lerdahl and Jackendoff's (1983) theory of grouping structure consists of 5 well-formedness rules, 7 preference rules and 2 other rules called 'transformational rules' that deal with cases where groups overlap.

2. Let's start by looking at the well-formedness rules.

3. GWFR 1 states that "any contiguous sequence of pitch-events, drum beats, or the like can constitute a group, and only contiguous sequences can constitute a group" (Lerdahl and Jackendoff, 1983, p. 37).

4. All this means is that a group must contain everything that happens in the music between two specified timepoints. It must not, for example, contain all the first notes in each bar or all the Ds in the piece. It is this rule that allows us to represent a group by using a slur or a phrase mark.

5. GWFR 2 specifies that "a piece constitutes a group" (Lerdahl and Jackendoff, 1983, p. 38). This rule is fairly self-evident. It simply specifies that a complete piece of music is always perceived as being a structural unit.

6. GWFR 3 states that "a group may contain smaller groups" (Lerdahl and Jackendoff, 1983, p. 38).

7. As we saw in this Mozart example, the whole melody is perceived as a group but it is also perceived as being constructed from two smaller, phrase-level groups, which can, in turn, be perceived to be constructed from three motive-like groups.

8. GWFR 4 states that "if a group G_1 contains part of a group G_2, then it must contain all of G_2" (Lerdahl and Jackendoff, 1983, p. 38).

9. All this means is that a group must not begin or end in the middle of some lower-level group or in the middle of an adjacent group at the same level. In other words, you never get grouping structures like this or this. This is intuitively fairly obvious: if this is a group [POINT TO A LOW LEVEL GROUP] and this is a group [POINT TO THE ADJACENT LOW LEVEL GROUP] then there cannot exist some higher level group that is perceived to end in the middle of this second group—if a higher level group boundary were perceived to exist here, then this lower-level group would also be perceived to end here.

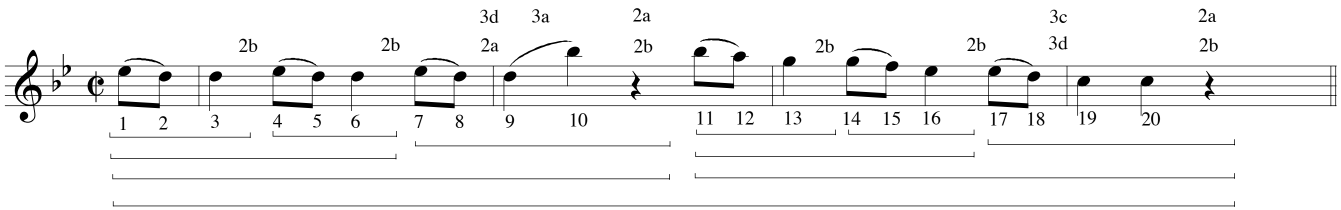
10. In fact, there are some special situations, which Lerdahl and Jackendoff (1983) call grouping overlaps or grouping elisions, where the last event in a group is simultaneously heard to be the first event in the next group. Such situations violate GWFR 4 and Lerdahl and Jackendoff (1983) propose two special, 'transformational rules' to deal with such cases.

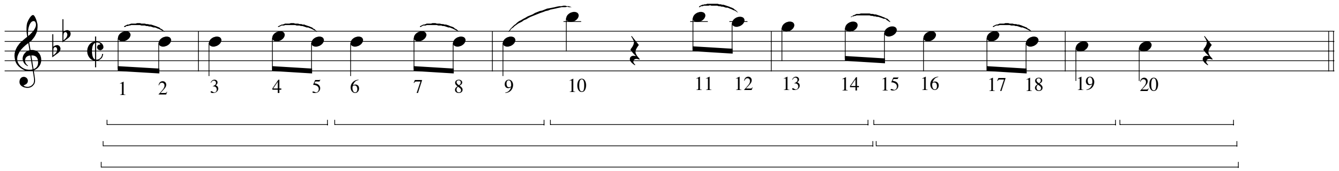
11. The fifth and final GWFR states that "if a group G_1 contains a smaller group G_2 then G_1 must be exhaustively partitioned into smaller groups" (Lerdahl and Jackendoff, 1983, p. 38).

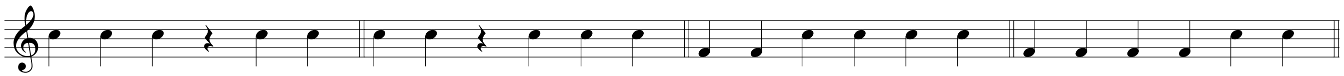
12. In other words, this sort of situation never happens: if these two segments are perceived to be groups and this segment is perceived to be a group at a higher level, then this segment here between the end of this lower-level group and the end of the higher-level group, will itself be perceived to be a group.

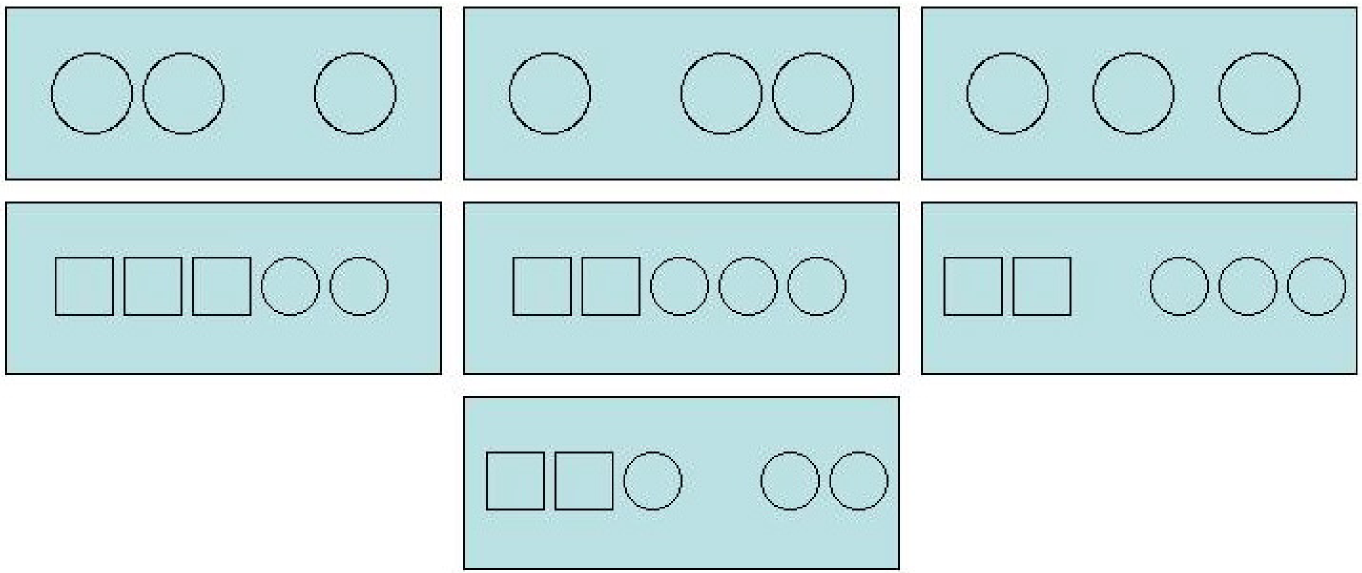
13. On the other hand, it's perfectly alright for some but not all of the groups at a particular level to be subdivided into subgroups, as happens, for example, here.

5. The Gestalt principles of proximity and similarity









5. The Gestalt principles of proximity and similarity

1. On their own, the grouping well-formedness rules cannot be used to predict the grouping structure that we hear when we listen to a passage of music.
2. The grouping well-formedness rules simply define a large set of grouping structures that (we hope) contains the grouping structure that we actually hear.
3. For example, this grouping structure here satisfies the grouping well-formedness rules but it certainly isn't the structure that most people would hear when they listen to this melody.
4. A good way to test a grouping structure is to see what it sounds like when you insert rests at the group boundaries. Here's what this predicted grouping structure sounds like when we do that [PLAY GOOD GROUPING]. And here's what this grouping structure sounds like [PLAY BAD GROUPING]. I think most of you will agree this grouping structure is not a good description of the way that we naturally interpret this melody.
5. So we need to find some principles that will allow us to predict the grouping structure that listeners actually hear.
6. As I mentioned before, it seems that the principles governing the way that we group events in a melody are closely related to those that govern the way that we group objects in a visual scene.
7. In the 1920s and 1930s, a group of German psychologists, in particular Wertheimer (1938), Köhler (1947) and Koffka (1935), developed a theory called the Gestalt theory of perception in which they proposed that the way we naturally group objects together in a visual scene is governed by a small number of simple principles.
8. Two of these principles in particular—the so-called principles of similarity and proximity—seem to be also involved in the way that we group events in a passage of music.
9. To illustrate these principles, I'd like you, please, [PICK SOMEONE IN THE AUDIENCE] to describe what you see here in this picture [TWO CIRCLES ON LEFT AND ONE ON RIGHT]. Now could you please describe what you see in this picture [TWO CIRCLES ON RIGHT AND ONE ON LEFT]. Now please describe what you see in this picture [THREE CIRCLES IN A ROW].
10. Notice that he described the first picture as consisting of two circles on the left and one on the right, the second picture as containing one circle on the left and two on the right and the third as containing three circles in a row.
11. This shows that in the first picture he automatically grouped the middle circle with the one on the left. However, when the middle circle was closer to the right-hand circle, he grouped it with that one. Finally, in this picture where the circles are an equal distance apart, he simply grouped all three circles together.

12. This illustrates the Gestalt principle of proximity which states that closer elements are grouped together in preference to those that are spaced further apart.

13. Now could you please describe what you see here [THREE SQUARES + TWO CIRCLES]. And here [TWO SQUARES + THREE CIRCLES].

14. Note that she did not describe this picture as consisting of two squares, another square and two circles, nor did she describe this picture as containing two squares and a circle and another two circles. In both cases, she grouped all the squares together and all the circles together.

15. This illustrates the Gestalt principle of similarity which states that we tend to group similar objects together.

16. Wertheimer (1938) himself pointed out that these two principles also seem to operate in music. For example, in this fragment here, we would tend to hear the third note as being grouped with the first two whereas, in this case, we would tend to hear the third note as being grouped with the last two notes.

17. This seems to be a case, then, where the principle of proximity is operating.

18. Similarly, in both these cases we tend to hear the Fs as belonging together and the Cs as belonging together in another group, suggesting that the principle of similarity may be operating here.

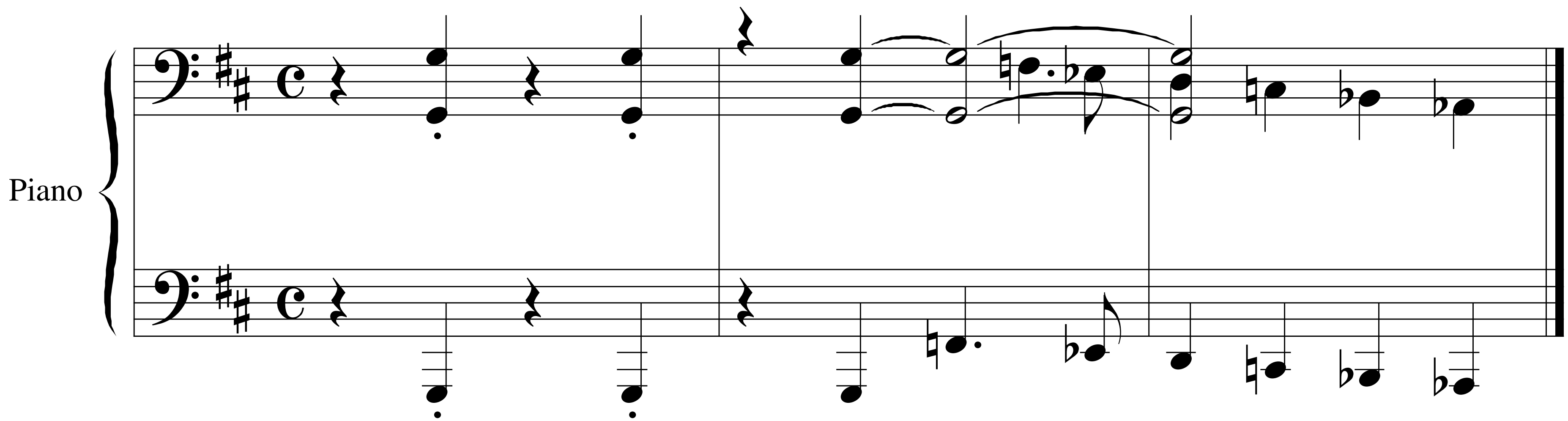
19. It's also possible for both of these principles to operate in a single visual or auditory field. For example, Gestalt theory would predict that we more definitely separate the circles from the squares in this picture than we do in this one because here both principles are operating at once and reinforcing each other.

20. However, in this example, the principle of proximity pushes us towards grouping the circle with the squares but the principle of similarity pushes us to group it with the other circles. The two principles conflict causing us to be less definite about our interpretation.

21. As we shall see, two of Lerdahl and Jackendoff's (1983) grouping preference rules are, in fact, simply reformulations of the Gestalt principles of proximity and similarity.

6. Lerdahl and Jackendoff's (1983) Local Detail Grouping Preference Rules:

GPR 1

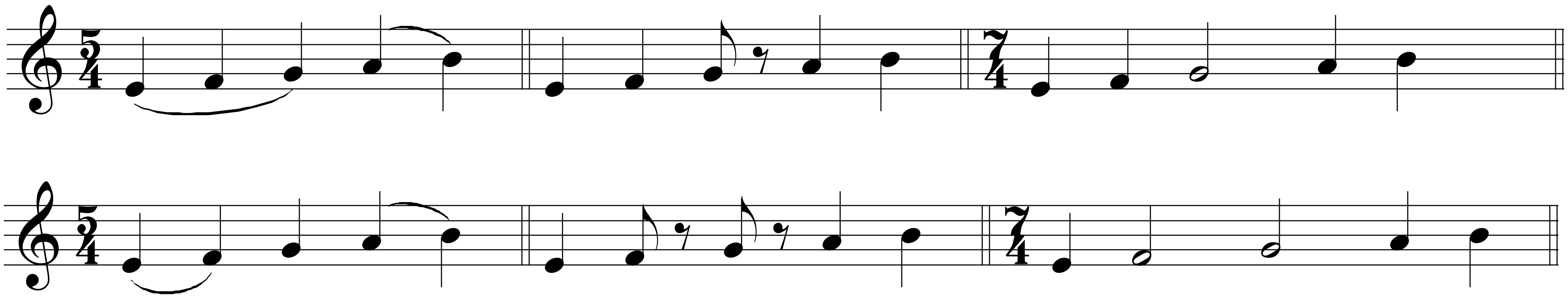


GPR 1, alternative form Avoid analyses with very small groups—the smaller, the less preferable.

6. Lerdahl and Jackendoff's (1983) Local Detail Grouping Preference Rules: GPR 1

1. Lerdahl and Jackendoff's (1983) theory of grouping structure contains seven grouping preference rules of which three, the so-called 'Local Detail Rules', deal with small groups and the other four deal with larger-level grouping.
2. The first grouping preference rule states that we tend to "avoid analyses with very small groups—the smaller, the less preferable" (Lerdahl and Jackendoff, 1983, p. 43).
3. Lerdahl and Jackendoff (1983) claim that a group can only contain a single note if it is either strongly isolated or if it functions as a motive all by itself.
4. I think the first two events in Liszt's Sonata in B minor probably count as rare examples of GPR 1 being overridden. [PLAY LISZT SONATA IN B MINOR].

7. Lerdahl and Jackendoff's (1983) Local Detail Grouping Preference Rules: GPR 2 (Proximity)



GPR 2 (Proximity) Consider a sequence of four notes n_1, n_2, n_3, n_4. All else being equal, the transition n_2–n_3 may be heard as a group boundary if

a. (Slur/Rest) the interval of time from the end of n_2 to the beginning of n_3 is greater than that from the end of n_1 to the beginning of n_2 and that from the end of n_3 to the beginning of n_4, or if

b. (Attack-Point) the interval of time between the attack points of n_2 and n_3 is greater than that between the attack points of n_1 and n_2 and that between the attack points of n_3 and n_4.

7. Lerdahl and Jackendoff's (1983) Local Detail Grouping Preference Rules: GPR 2 (Proximity)

1. The second grouping preference rule is Lerdahl and Jackendoff's (1983) version of the Gestalt rule of proximity. It states that if we have "a sequence of four notes n_1, n_2, n_3, n_4 [then,] [a]ll else being equal, the transition n_2–n_3 may be heard as a group boundary if

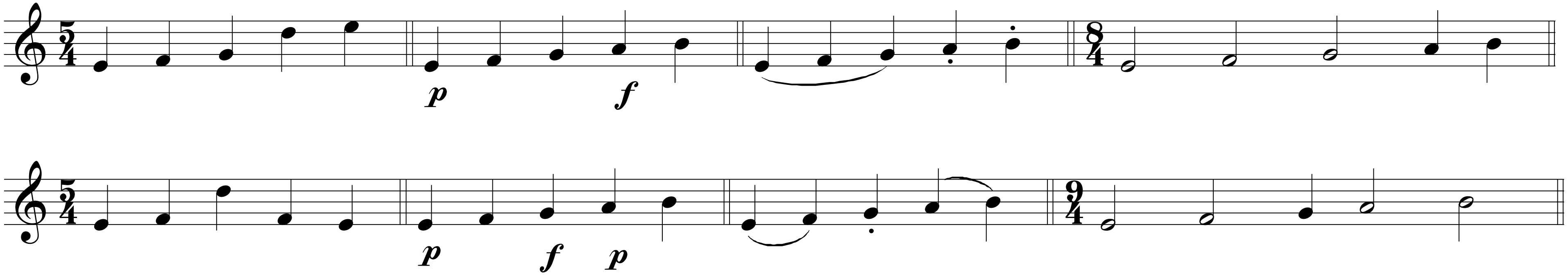
   a. (Slur/Rest) the interval of time from the end of n_2 to the beginning of n_3 is greater than that from the end of n_1 to the beginning of n_2 and that from the end of n_3 to the beginning of n_4, or if

   b. (Attack-Point) the interval of time between the attack points of n_2 and n_3 is greater than that between the attack points of n_1 and n_2 and that between the attack points of n_3 and n_4."

2. For example, GPR 2a predicts there will be a group boundary here and here and GPR 2b predicts there will be a boundary here.

3. However, GPR 2 does not predict any boundaries in these examples [EXPLAIN on SLIDE].

8. Lerdahl and Jackendoff's (1983) Local Detail Grouping Preference Rules: GPR 3 (Change)



GPR 3 (Change) Consider a sequence of four notes n_1, n_2, n_3, n_4. All else being equal, the transition n_2–n_3 may be heard as a group boundary if

a. (Register) the transition n_2–n_3 involves a greater intervallic distance than both n_1–n_2 and n_3–n_4, or if

b. (Dynamics) the transition n_2–n_3 involves a change in dynamics and n_1–n_2 and n_3–n_4 do not, or if

c. (Articulation) the transition n_2–n_3 involves a change in articulation and n_1–n_2 and n_3–n_4 do not, or if

d. (Length) n_2 and n_3 are of different lengths and both pairs n_1, n_2 and n_3, n_4 do not differ in length."

(One might add further cases to deal with such things as change in timbre or instrumentation.)

8. Lerdahl and Jackendoff's (1983) Local Detail Grouping Preference Rules: GPR 3 (Change)

1. The third grouping preference rule is Lerdahl and Jackendoff's (1983) version of the Gestalt rule of similarity. It states that if we have "a sequence of four notes n_1, n_2, n_3, n_4 [then] [a]ll else being equal, the transition n_2–n_3 may be heard as a group boundary if

   a. (Register) the transition n_2–n_3 involves a greater intervallic distance than both n_1–n_2 and n_3–n_4, or if

   b. (Dynamics) the transition n_2–n_3 involves a change in dynamics and n_1–n_2 and n_3–n_4 do not, or if

   c. (Articulation) the transition n_2–n_3 involves a change in articulation and n_1–n_2 and n_3–n_4 do not, or if

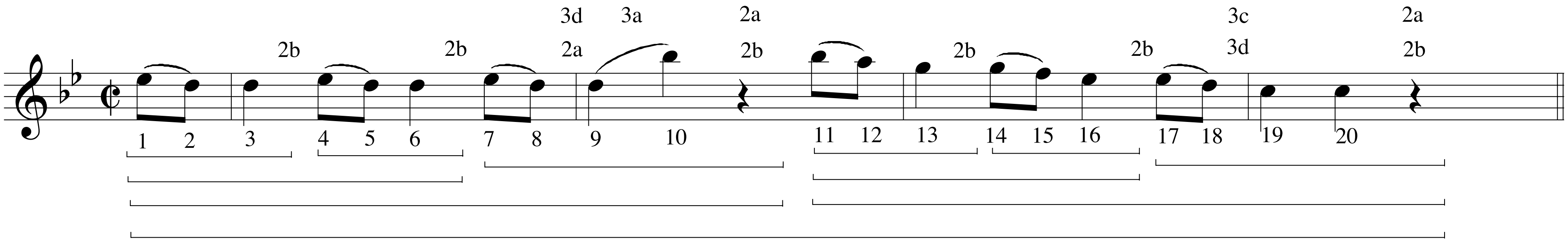
   d. (Length) n_2 and n_3 are of different lengths and both pairs n_1, n_2 and n_3, n_4 do not differ in length."

   [They add that] "further cases [may be added] to deal with such things as change in timbre or instrumentation" (Lerdahl and Jackendoff, 1983, p. 46).

2. So, for example, GPR3a predicts a boundary here, GPR3b predicts a boundary here, GPR3c predicts a boundary here and GPR3d predicts a boundary here.

3. However GPR 3 predicts no boundaries in any of these examples because in each case, two changes occur consecutively.

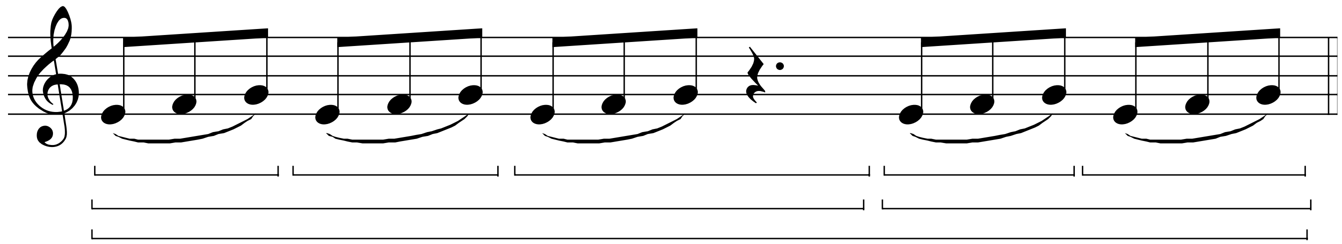
9. Example application of the local detail rules

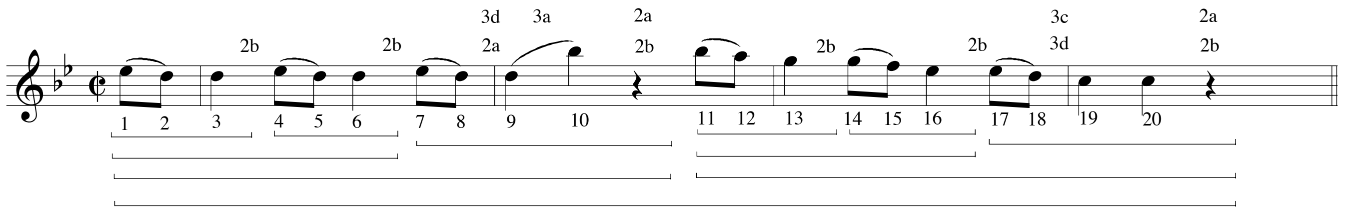


9. Example application of the local detail rules

1. This figure here shows what happens when we apply the first three local detail rules to the opening melody of the Mozart G minor symphony.
2. Underneath the staff we have the intuitively reasonable grouping structure that I showed you earlier on and the labels above the staff show all the applications of GPRs 2 and 3.
3. As you can see, all of the places where these rules apply are indeed perceived group boundaries apart from in three cases: the transition from note 8 to note 9, the transition from note 9 to note 10 and the transition from note 18 to note 19.
4. Lerdahl and Jackendoff (1983, p. 48) claim that we do not hear a boundary between notes 9 and 10 because the boundary between notes 10 and 11 is more salient and if we had boundaries in both places, this would result in a group containing only one event which would violate GPR 1.
5. Their explanation for why we don't hear boundaries between notes 8 and 9 and between notes 18 and 19 involves reference to a rule that we haven't yet discussed so I'll come back to that in a few minutes.

10. Lerdahl and Jackendoff's (1983) Larger-Level Grouping Rules: GPR 4 (Intensification)



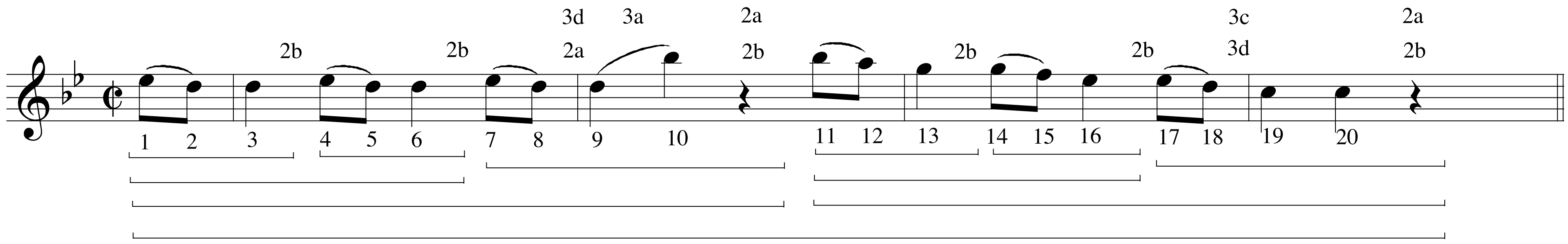


GPR 4 (Intensification) Where the effects picked out by GPRs 2 and 3 are relatively more pronounced, a larger-level group boundary may be placed.

10. Lerdahl and Jackendoff's (1983) Larger-Level Grouping Rules: GPR 4 (Intensification)

1. Lerdahl and Jackendoff (1983) propose that our perception of higher-level grouping structure is governed by four preference rules, of which the first, GPR 4, states that "where the effects picked out by GPRs 2 and 3 are relatively more pronounced, a larger-level group boundary may be placed" (Lerdahl and Jackendoff, 1983, p. 50).
2. They provide this example here which illustrates how GPR 4 can be used to predict higher-level grouping boundaries.
3. In this example, GPRs 2a and 2b predict group boundaries between each group of slurred triplets. However the offset-to-onset interval between the end of the third group of triplets and the beginning of the fourth set is much longer than that between the other triplet groups. Therefore GPR4 predicts that there will be a larger-level group boundary between the third and fourth sets of triplets.
4. Note that GPR4 also predicts that there will be a higher-level grouping boundary between the two phrases in the Mozart G minor Symphony melody here because of the extra long gap between notes 10 and 11.

11. Lerdahl and Jackendoff's (1983) Larger-Level Grouping Rules: GPR 5 (Symmetry)

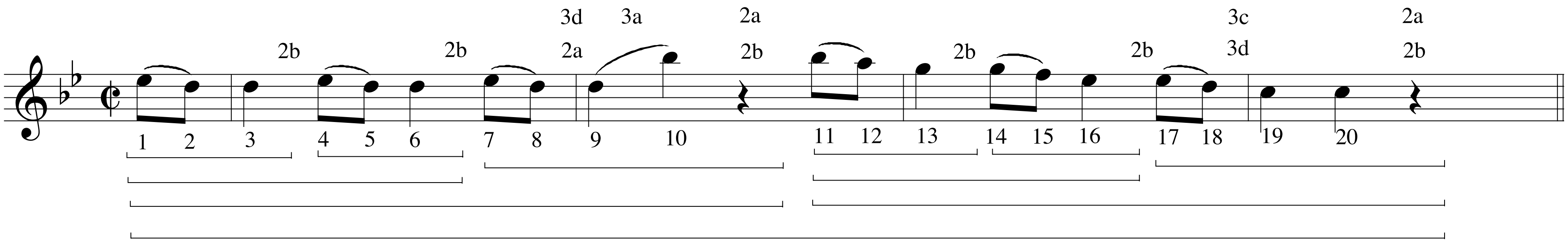


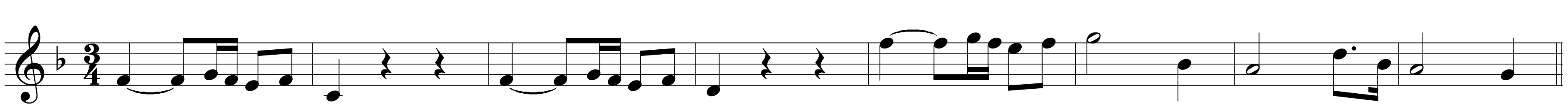
GPR 5 (Symmetry) Prefer grouping analyses that most closely approach the ideal subdivision of groups into two parts of equal length.

11. Lerdahl and Jackendoff's (1983) Larger-Level Grouping Rules: GPR 5 (Symmetry)

1. The second larger-level grouping rule, GPR 5, predicts that we "prefer grouping analyses that most closely approach the ideal subdivision of groups into two parts of equal length" (Lerdahl and Jackendoff, 1983, p. 49).
2. In this Mozart example, GPR 5 applies here, between notes 10 and 11, and serves to reinforce GPR 4s effect of dividing the melody into two halves.
3. It also operates here between notes 6 and 7 and here between notes 16 and 17 where it causes these two low-level groups (11-13,14-16) to be grouped together into an intermediate-level group that balances the group 17–20.
4. Lerdahl and Jackendoff (1983) claim that it is our desire for symmetrical grouping divisions that gives rise to our perception of ternary groupings as being less stable than binary ones.
5. However, I don't know of any clear evidence that we *do* find ternary groupings less stable, so I personally don't find this argument very convincing.
6. Here, for example, I am quite happy to hear this first phrase as being composed from three subgroups and I'm not sure that I really hear the middle subgroup as being more closely related to the first subgroup than to the third.

12. Lerdahl and Jackendoff's (1983) Larger-level Grouping Rules: GPR 6 (Parallelism)



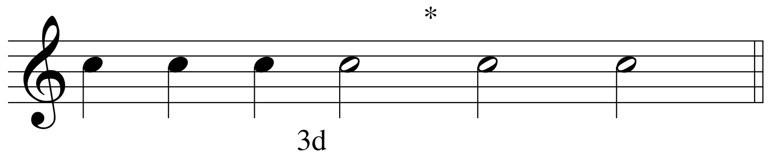


GPR 6 (Parallelism) Where two or more segments of the music can be construed as parallel, they preferably form parallel parts of groups.

12. Lerdahl and Jackendoff's (1983) Larger-level Grouping Rules: GPR 6 (Parallelism)

1. Their sixth grouping preference rule states that "where two or more segments of the music can be construed as parallel, they preferably form parallel parts of groups" (Lerdahl and Jackendoff, 1983, p. 51).
2. Lerdahl and Jackendoff (1983) don't tell us exactly how we're supposed to decide whether two segments are parallel but basically it seems that we're supposed to "construe two groups as being parallel' if they are similar in some way—for example, if they have the same rhythm or the same pitch contour or the same harmonic progression.
3. Lerdahl and Jackendoff (1983) stress that the rule does not state that parallel segments must form entire groups but only that they should form parallel *parts* of groups.
4. The rule is expressed in this way in order to account for the commonly occurring situation where two groups (e.g., an antecedent and consequent phrase) begin in the same way but end differently.
5. This happens, for example, in this melody from the beginning of Beethoven's Quartet Op.18, No.1.
6. The first bar of the opening phrase is identical to bar 3 which is the first bar of the second phrase but neither of these two bars forms a group. Also, bar 5 is a transposition of bar 1 and is therefore parallel to it.
7. Because the first bar begins with a group boundary, the parallelism rule predicts that there will be group boundaries at the beginning of bars 3 and 5 also.
8. In the Mozart melody, GPR 6 reinforces GPRs 4 and 5 to predict a strong group boundary between notes 10 and 11.
9. On a lower level of structure, it predicts that segments 1–3, 4–6 and 7–9 will form parallel parts of groups. Therefore, because no grouping boundary is heard between notes 2 and 3, and between 5 and 6, GPR 6 predicts that no boundary will be heard between 8 and 9 either.
10. The seventh GPR concerns the interaction between grouping structure and the time-span reduction and since I'm not considering time-span reduction, I won't talk about that particular rule.
11. In addition there are two transformational rules that deal with cases where groups are perceived to overlap each other. But I won't talk about them either right now.

13. Deliège's (1987) Experiments on Lerdahl and Jackendoff's (1983) grouping theory: Experiment 1



- Both musicians and non-musicians gave results that were significantly in accordance with the rules;
- Musicians responded significantly more in accordance with the rules than non-musicians;
- Non-musicians did not need significantly more repetitions than musicians;
- Both types of subject needed significantly more repetitions on the trials; when they did not respond in accordance with the rules;
- Most commonly used rules: GPR 2b (Attack-Point), GPR 3b (Change in dynamics) and GPR 3 timbre rule.
- Musicians and non-musicians use essentially the same strategy to group melodies.
- Subjects use a modified version of GPR 2b rather than GPR 3d.

13. Deliège's (1987) Experiments on Lerdahl and Jackendoff's (1983) grouping theory: Experiment 1

1. Four years after the publication of GTTM, Irène Deliège published a paper describing two experiments that she had carried out in order to investigate how well Lerdahl and Jackendoff's (1983) grouping theory accounted for the grouping structures perceived by experienced and inexperienced listeners.
2. In the first experiment, the subjects listened to short passages from instrumental or orchestral works by various composers ranging from Bach to Stravinsky and then indicated on an answer sheet how they perceived each passage to be divided up into subgroups.
3. The subjects were allowed to hear the passages as many times as they liked but the experimenter made note of how many times each subject heard each passage.
4. Deliège (1987) then compared the segmentations produced by the subjects with those predicted by GPRs 2 and 3—that is, Lerdahl and Jackendoff's (1983) versions of the Gestalt principles of proximity and similarity.
5. Deliège (1987, p. 337) found that
   (a) both musicians and non-musicians gave results that were significantly in accordance with the rules;
   (b) musicians responded significantly more in accordance with the rules than non-musicians;
   (c) non-musicians did not need significantly more repetitions than musicians; and
   (d) both types of subject needed significantly more repetitions on the trials when they did not respond in accordance with the rules.
6. Recall that Lerdahl and Jackendoff (1983, p. 46) proposed that it might be useful to add a subrule to GPR3 that predicts that listeners will tend to hear a group boundary when there is a change in timbre.
7. Deliège (1987) found that there were three rules that all subjects used particularly often. These were
   (a) GPR 2b, the Attack-Point rule;
   (b) GPR 3b, the Change in dynamics rule; and
   (c) a version of GPR 3 that predicts a grouping boundary when there is a change in timbre.
8. She found that the other rules were used less by everyone, but *much* less by the non-musicians.
9. What these results seem to show is that both musicians and non-musicians use basically the same mechanisms to infer grouping structure from musical surfaces: essentially, they use the relative lengths of events and changes in loudness and timbre.
10. However, musical training seems to make listeners more sensitive to additional cues such as changes in articulation and relatively large pitch intervals.

11. Recall that GPR 3 d predicts that when a listener is presented with a sequence of short notes followed by a sequence of long notes, he or she will tend to hear a group boundary at the beginning of the first long note (see figure).

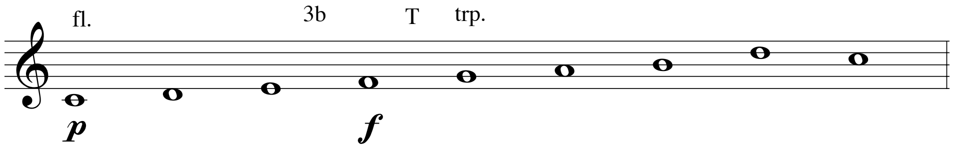
12. However, a listener only knows that the long note is the first of a sequence of long notes when he or she hears the *third* long note (or, at the very earliest, some way through the second long note).

13. GPR 3d therefore implies that when the listener realises that the long note is the first in a sequence of long notes some time in the middle of the second long note in the sequence, he or she will retrospectively hear a grouping boundary before the first long note.

14. In her experiment, Deliège (1987) found that, in practice, in this kind of situation, listeners actually tend to hear a group boundary between the first and second long notes in the sequence.

15. In other words, they seem to be using a modified version of the Attack-Point rule, GPR 2b, which operates even when the longer interval is not followed by a shorter time interval as demanded by Lerdahl and Jackendoff (1983).

14. Deliège's (1987) Experiments on Lerdahl and Jackendoff's (1983) grouping theory: Experiment 2

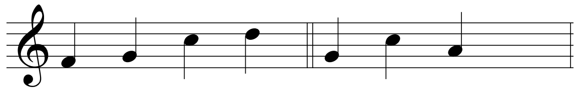


- Lerdahl and Jackendoff (1983) do not specify means for measuring strength of application of a given rule in a given situation.
- Deliège (1987) investigated rule conflicts.
- 93% non-musicians, 97% musicians gave responses in accordance with rules.
- Change in timbre was most popularly used cue for both musicians and non-musicians.

14. Deliège's (1987) Experiments on Lerdahl and Jackendoff's (1983) grouping theory: Experiment 2

1. One of the problems with Lerdahl and Jackendoff's (1983) grouping theory is that they do not specify how to measure the strength of the effect of any particular rule application.
2. In her second experiment, Deliège (1987) presented subjects with artificially constructed melodic phrases based on rising or falling scales, designed to see what happens when the grouping preference rules conflict with each other.
3. Here, for example, is one of the patterns she used in her experiment. In this case, GPR 3b (the dynamics rule) predicts a boundary between notes 3 and 4 but the change in timbre between notes 4 and 5 predicts a boundary here.
4. Listeners were told that in each trial they were only permitted to put one grouping boundary.
5. In this experiment, Deliège (1987) found that 93% of non-musicians and 97% of musicians gave responses that were in accordance with the rules. (That is, in the vast majority of cases, subjects did choose one of the two predicted boundaries in each trial.)
6. A statistical analysis showed that musicians responded in accordance with the rules significantly more often than non-musicians.
7. Recall that in the first experiment, she found that subjects most often used the attack-point rule, change in dynamics and change in timbre to as grouping cues. In this second experiment, change in timbre was the most popular rule for both musicians and non-musicians.
8. Deliège (1987) also found that listeners hardly ever use change in contour as a grouping cue.
9. Finally, she found that non-musicians behaved more in accordance with the predictions of the rules in the second experiment where they were presented with short, artificially constructed fragments than they did when presented with extracts from "real" orchestral and instrumental music.

15. Cambouropoulos's (1998) Local Boundary Detection Model



- *General Computational Theory of Musical Structure.*
- Generates segmentation and then clusters segments by similarity.
- *General Identity-Change Rule* Grouping boundaries may be introduced only between two different entities. Identical entities do not suggest any boundaries between them.
- *Identity-Change Rule* Amongst three successive objects, boundaries may be introduced on either of the consecutive intervals formed by the objects if these intervals are different. If both intervals are identical, no boundary is suggested.
- *Proximity Rule* Amongst three successive objects that form different intervals between them, a boundary may be introduced on the larger interval, i.e., those two objects will tend to form a group that are closer together (or more similar to each other).

15. Cambouropoulos's (1998) Local Boundary Detection Model

1. I'm going to finish off this lecture by introducing you to two recent computational models of musical grouping, namely, those of Cambouropoulos (1998) and Temperley (2001).

2. In his Ph.D. thesis which he completed in 1998, Cambouropoulos (1998) presents what he calls a "General Computational Theory of Musical Structure" that predicts the lower-level grouping structure that will be heard by a listener and then clusters the various groups into categories, each category containing a set of parallel groups.

3. Cambouropoulos (1998, p. 67) points out that the Gestalt principles of proximity and similarity can be seen as being two sides of the same coin.

4. Consider, for example, this sequence of four notes here. Lerdahl and Jackendoff's (1983) GPR 3a predicts that we hear a grouping boundary between the second and third note because the interval between them is larger than that between the first and second note and that between the third and fourth note.

5. Cambouropoulos (1998, p. 67) points out that this situation can be seen either from the viewpoint of the similarity principle or from that of the proximity principle:

   (a) One could say that the second pitch is more similar to the first pitch than it is to the third pitch and the third pitch is more similar to the fourth pitch than it is to the second pitch, thus suggesting that it is the principle of similarity that causes us to hear a boundary between the second and third pitches.

   (b) Alternatively, one could say that the second pitch is *closer* to the first pitch within a particular pitch space, than it is to the third pitch and that the third pitch is *closer* to the fourth pitch than it is to the second pitch, suggesting that it is the principle of proximity that causes us to hear a boundary between the second and third pitches.

6. Instead of using the Gestalt rule of similarity, Cambouropoulos (1998, p. 68) introduces a more elementary rule which he calls the "General Identity-Change Rule". This rule states quite simply that "grouping boundaries may be introduced only between two different entities. Identical entities do not suggest any boundaries between them."

7. To derive a grouping structure, Cambouropoulos (1998, p. 68) uses a more specific version of this rule which can be applied to sequences of *intervals* (e.g., pitch intervals, inter-onset intervals and so on). This, more specific version of the rule goes like this: "Amongst three successive objects, boundaries may be introduced on either of the consecutive intervals formed by the objects if these intervals are different. If both intervals are identical, no boundary is suggested."

8. For example, let's imagine that the three objects are these three notes (G4,C5,A4) and we are considering the chromatic pitch intervals between them. The interval from the G4 to the C5 is different from that from C5 to A4 therefore, according to Cambouropoulos's (1998) Identity-Change Rule, a boundary may be introduced either between the G4 and the C5 or between the C5 and the A4.

9. Clearly, in such cases we do not in general want a boundary on both of the consecutive positions. To avoid this, Cambouropoulos (1998, p. 69) uses a version of the proximity rule which he states as follows: "Amongst three successive objects that form different intervals between them, a boundary may be introduced on the larger interval, i.e., those two objects will tend to form a group that are closer together (or more similar to each other)."

10. In our previous example here, this, 'Proximity Rule' predicts that the boundary will actually fall between the G4 and the C5.

11. Cambouropoulos (1998, p. 70) applies weighted versions of these rules to various different intervallic representations of melodic sequences to obtain a measure of boundary strength for every interval. The intervals that have locally maximum boundary strengths are then the predicted group boundaries.

12. Cambouropoulos (1998, p. 70) claims that Lerdahl and Jackendoff's (1983) local detail grouping preference rules form a special case of his theory.

16. Temperley's (2001) computational model of melodic phrase structure

PSPR 1 (Gap rule) Prefer to locate phrase boundaries at a) large inter-onset-intervals and b) large offset-to-onset intervals.

PSPR 2 (Phrase Length Rule) Prefer phrases to have roughly eight notes.

PSPR 3 (Metrical Parallelism Rule) Prefer to begin successive groups at parallel points in the metrical structure.

16. Temperley's (2001) computational model of melodic phrase structure

1. Finally, I'll just briefly describe David Temperley's (2001) computational model of melodic phrase structure.
2. Unlike Cambouropoulos's (1998) model which, at least in principle, is capable of generating several different levels of grouping structure, Temperley's (2001) system only attempts to find one level of grouping—namely, the one that corresponds to the perceived phrase structure.
3. Like Cambouropoulos's (1998) model, it can only be used to analyse melodies.
4. Like his model of metrical structure which I briefly described last week, Temperley's (2001) model of melodic phrase structure takes the form of a computer program that accepts a MIDI-like representation of the melody as input and generates as output a prediction of the phrase-structure that will be heard by a listener.
5. Also like his model of metrical structure, Temperley's (2001) grouping structure program is an implementation of a system of preference rules. Temperley (2001) uses the optimisation technique of dynamic programming to find efficiently a grouping structure that best satisfies the three preference rules in his model.
6. These three preference rules are shown here:
7. The first rule proposes that we "prefer to locate phrase boundaries at a) large inter-onset-intervals and b) large offset-to-onset intervals" (Temperley, 2001, pp. 68, 358).
8. This rule is simply a modification of Lerdahl and Jackendoff's (1983) GPR 2. It is a simple variation on the Gestalt principle of proximity.
9. The second rule (Temperley, 2001, pp. 69, 358) is an astonishingly simplistic, ad hoc rule that claims we prefer phrases to contain roughly 8 notes.
10. In practice, the average number of notes per phrase depends on the style, genre and instrumentation of the music.
11. For example, as Temperley (2001, p. 82) himself points out, in instrumental music, phrases routinely contain many more than 8 notes. Even in the sample of simple folk songs that Temperley used to test his program, the average length of phrases turned out to be 9.2 notes, which meant that the model performed best on that particular corpus when this second rule was modified to favour 10-note phrases (Temperley, 2001, p. 74).
12. This illustrates just how ad hoc this second rule is.
13. Temperley's third rule is an attempt to model the effect of parallelism on grouping.
14. Temperley acknowledges that identifying parallelism (i.e., significant repetitions) is an important factor in grouping structure (Temperley, 2001, p. 69). However, he considers it to be a huge and complex problem and so doesn't attempt to solve it in his preference rule systems, even though many of the mistakes made by these systems seem to be due to their failure to take parallelism into account.

15. Temperley tested his theory by running it on a corpus of 65 simple folk songs containing 257 phrase boundaries. His program failed to identify 63 of the boundaries in the database and predicted 66 boundaries where none occurred in the database.

References

Cambouropoulos, E. (1998). *Towards a General Computational Theory of Musical Structure*. Ph.D. thesis, University of Edinburgh.

Deliège, I. (1987). Grouping conditions in listening to music: An approach to Lerdahl and Jackendoff's Grouping Preference Rules. *Music Perception*, **4**(4), 325–360. Cited by Krumhansl (1990).

Koffka, K. (1935). *Principles of Gestalt Psychology*. Harcourt, Brace and World, New York. Cited by Krumhansl (1990).

Köhler, W. (1947). *Gestalt Psychology*. Liveright, New York. Cited by Krumhansl (1990).

Krumhansl, C. L. (1990). *Cognitive Foundations of Musical Pitch*, volume 17 of *Oxford Psychology Series*. Oxford University Press, New York and Oxford.

Lerdahl, F. and Jackendoff, R. (1983). *A Generative Theory of Tonal Music*. MIT Press, Cambridge, MA.

Temperley, D. (2001). *The Cognition of Basic Musical Structures*. MIT Press, Cambridge, MA.

Wertheimer, M. (1938). Laws of organization in perceptual forms. In W. D. Ellis, editor, *A Source Book of Gestalt Psychology*. Routledge and Keegan Paul, London. Originally published in 1923.