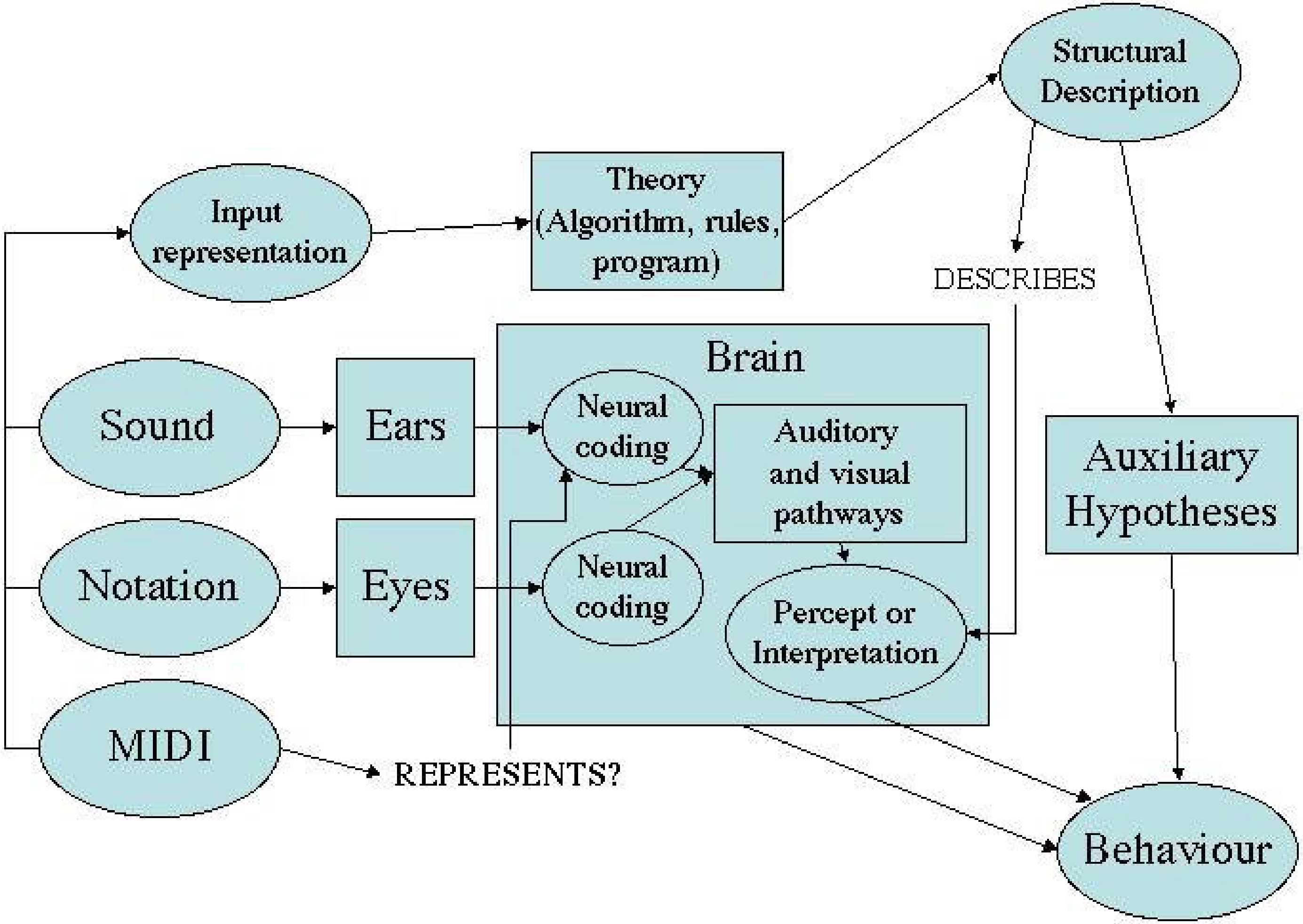
Musical Metre

David Meredith
Department of Computing,
City University, London.
`dave@titanmusic.com`
`www.titanmusic.com`

BMus/BSc in Music, Part Two Module: Music Perception and Cognition
Department of Music, City University, London
Tuesday, 13 May 2003.

1. Theories of music perception and cognition



1. Theories of music perception and cognition

1. In the next few lectures I'm going to focus on the problem of constructing and testing formal models and computational theories of music perception and cognition.

2. Most recent theories of music perception and cognition have taken the form of a system of rules or an algorithm or a computer program that takes a representation of a musical passage as input and generates as output a structure that is supposed to describe certain aspects of how the passage is perceived or interpreted by a listener.

3. For example, a theory of metre perception might take the form of a computer program that analyses a digital audio file containing a recording of a passage of music and predicts where the barlines would occur in a correctly notated score of the passage.

4. The theories that I'm going to describe vary significantly in the type of music representations that they accept as input:

    (a) some of the models process digital audio signals representing the sound that enters the listener's ear;

    (b) some models require information that is not explicitly available in the sound but is typically given in a notated score (e.g., barlines or pitch names); and

    (c) other models assume that the input gives information about the actions that need to be carried out in order to produce a performance of the passage represented—for example, a MIDI file or a piano-roll-like representation. It's even been suggested that a MIDI file is a good approximation to how the sound is encoded in the auditory nerve.

5. The choice of input representation should depend on precisely what the model is intended to be a model *of*. For example, if your intention is to construct a model of how an expert music analyst derives an analysis of a piece by studying a score of the piece, then it would be appropriate to use some kind of score representation as input. But if your intention is to model the mental processes involved when a listener interprets a passage as he or she listens to it, then using a score representation might not be appropriate because a score contains much information that is not explicitly available to the listener in an audio signal.

6. As I've already said, most of the theories I'm going to consider analyse a passage of music and generate some kind of structure that is intended to describe certain aspects of how the passage is perceived or interpreted by a listener. In other words, the structural descriptions generated by the theories are supposed to correctly represent what a listener feels or thinks while or after listening to the passages of music given as input to the theories.

7. However, to be useful and testable, a theory must do more than just predict what people will *think* or *feel* in response to some stimulus.

8. This is because thinking and feeling go on inside people's heads and consequently cannot be directly observed. We can, of course, ask subjects to describe through introspection what they are thinking and feeling but there are two main reasons why this is usually not a good strategy to adopt:
   (a) it's impossible to independently verify that the subject's report is a good description of what he or she is thinking or feeling;
   (b) subjects are often either very bad at describing their own mental processes and states of mind or incapable of doing so.

9. So, to be useful and testable, it must be possible to use a psychological theory to make predictions about what people will *do* in response to some stimulus—that is, how they will *behave* in certain types of circumstances. This usually means that the structural descriptions have to be supplemented by 'auxiliary hypotheses' that propose feasible mechanisms by which a given structural description might cause certain observed behaviour.

10. Now you might suggest that we can use PET or MRI scans of the brain to "see what people are thinking and feeling". But such a scan doesn't tell us what someone's thinking—it only tells us something about what is physically going on in the brain while someone is involved in some activity.

2. Theories of musical metre perception

What sorts of behaviour should a theory of metre be able to explain?

- times when listeners tap their feet when listening to music
- periodicities of dance movements
- notated time-signatures and barlines in scores
- how difficult rhythms are to learn (Povel and Essens, 1985)

2. Theories of musical metre perception

1. In this lecture I'm going to focus on theories of musical metre perception.
2. What aspects of what kinds of behaviour should a theory of musical metre perception be able to make predictions about?
   (a) It should be able to predict the precise instants at which listeners tap their feet or fingers or clap their hands while listening to any passage of music.
   (b) It should be able to predict the periodicities of dance movements and the swaying movements that musicians make while they perform.
   (c) It should be able to predict where a student, composer or transcriber will choose to place the bar-lines when he or she has to transcribe a passage from a recording or write down a passage that he or she has just imagined.
   (d) Theories of metre perception have also been used to predict how difficult various rhythms are to remember (Povel and Essens, 1985).
3. Let's now do a little experiment that should illustrate a couple of important points about metre perception and psychological theories in general.
   (a) You're going to hear a melody and I want you to tap in time to it with the index finger of your dominant hand and I want you to do it in such a way that no-one around you can either hear or see you tapping.
   (b) You'll hear a melody repeated three times. The first two times it will be played unaccompanied but on the third repetition it will be accompanied by a regular beat played on a drum. I want you to listen to the melody and, as soon as you can, I want you to start tapping along in time to it silently with the index finger of your dominant hand. I want you tap with a regular beat so that the time interval between each pair of consecutive taps is constant. I want you tap at the most natural speed. And when you get to the third repetition, I want you to notice if the drum is beating at the same times that you are tapping at. Does anyone not understand what they have to do?
   (c) *PLAY metre_tactus.mid*

(d) How many of you found that the drum was not beating at the same points that you were tapping at? How many of you tapped faster than the drum beat? How many of you tapped more slowly than the drum beat? How many of you were tapping at roughly the same rate but at different time points?

(e) Now this time I want you to again tap a regular beat in time with the melody but I want you to tap faster than comes naturally.

(f) *PLAY metre_subtactus.mid.*

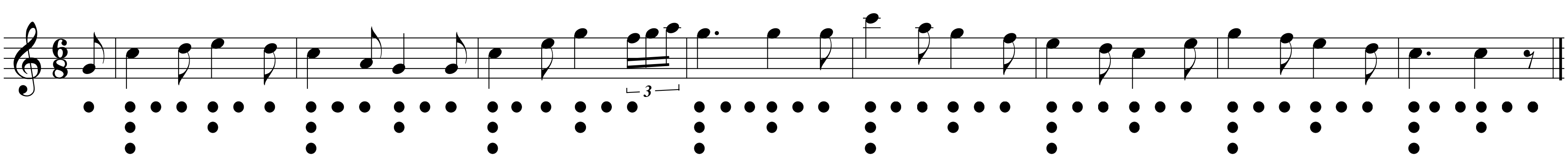
(g) How many of you found that the drum was not beating at the same points that you were tapping at? How many of you tapped faster than the drum beat? How many of you tapped more slowly than the drum beat? How many of you were tapping at roughly the same rate but at different time points?

(h) Now do the experiment for the final time, tapping more slowly than you did the first time (i.e., more slowly than what you consider to be the 'most natural speed').

(i) *PLAY metre_supertactus.mid.*

(j) How many of you found that the drum was not beating at the same points that you were tapping at? How many of you tapped faster than the drum beat? How many of you tapped more slowly than the drum beat? How many of you were tapping at roughly the same rate but at different time points?

3. Tapping experiment



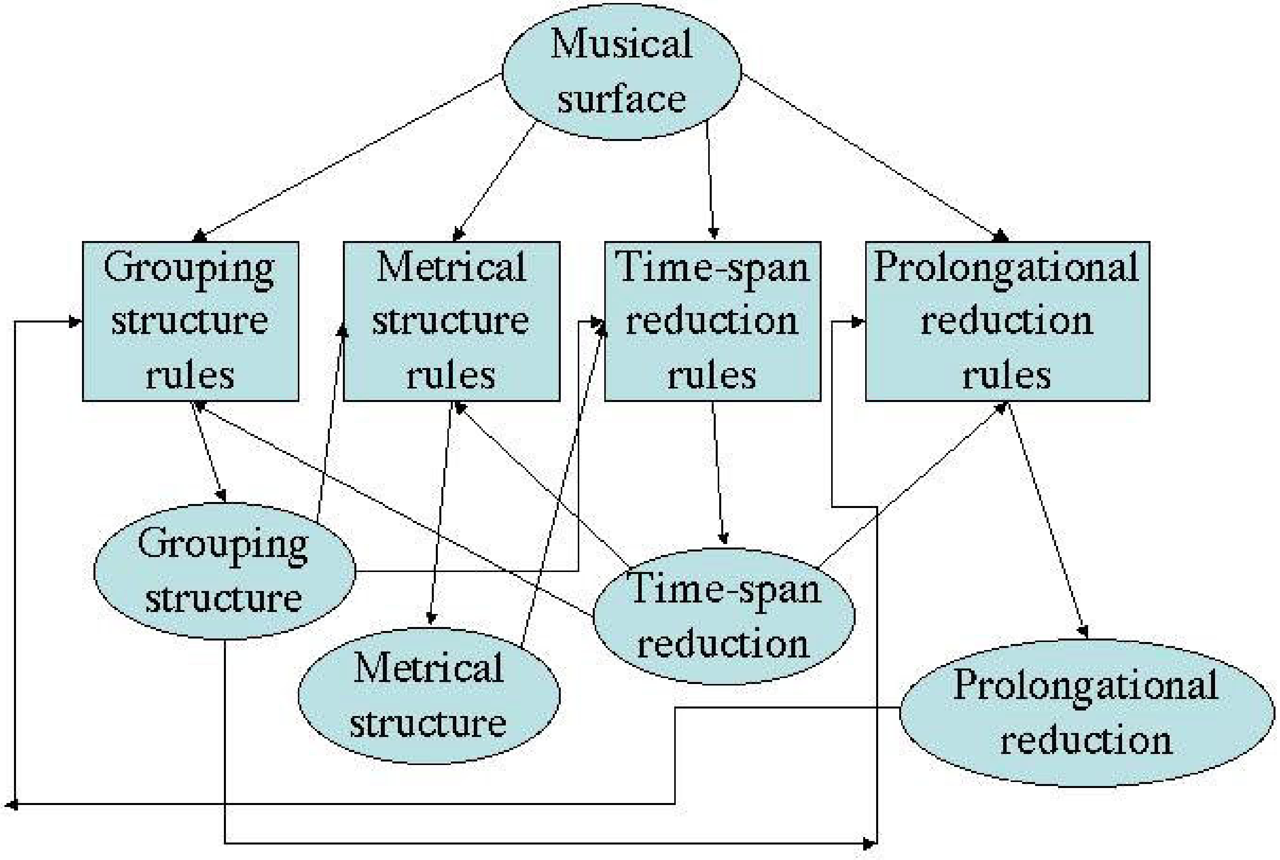
- psychological theories must account for the fact that different subjects give different responses
- general agreement about most appropriate times to tap or clap
- most listeners unconscious of mental processes involved in deciding when to tap

3. Tapping experiment

1. As I said, this experiment illustrates a number of important points:
2. First, it illustrates the fact that for many passages of music there is a 'most natural speed' at which one can tap in time. This most natural speed is called the tactus. I've represented the tactus here by placing dots under the events that you tapped at when you were tapping 'at the most natural rate'.
3. Second, it illustrates the fact that there are, in general, several different speeds at which one can tap in time to a passage of music—some faster than the tactus and some slower. Each of these speeds of tapping is called a 'metrical level' and it can be represented as a row of dots under the score as shown here. [SHOW ON SLIDE]
4. Moreover, when you tap at a speed that is slower than the tactus, you just miss out some of the tactus timepoints and when you tap at a speed that is faster than the tactus, you tap at all the tactus points, plus a few extra timepoints in between.
5. This suggests that the timepoints at which you tap when you're tapping more slowly than the tactus are more important or 'metrically stronger' than the tactus points that you miss out. Similarly, it suggests that the tactus timepoints are metrically stronger than all those extra timepoints that you tap at when you tap faster than the tactus.
6. Note that in this representation, the relative strength of a timepoint is represented by the number of dots underneath it.
7. The 'metrical structure' of a passage is simply a specification of the perceived metrical strength of each timepoint in the passage. So these sequences of dots drawn underneath the staff together form a representation of the metrical structure of the passage.
8. [IF DIFFERENT SUBJECTS TAPPED AT DIFFERENT RATES] then it showed that different subjects give different responses even when asked to perform the simplest of tasks. Psychological theories have to be able to account for this kind of thing by not only predicting which responses are most likely but also predicting that certain other responses are possible but less likely.

9. [IF MOST OF THE SUBJECTS TAPPED AT THE SAME RATE AS EACH OTHER AND THE DRUM] then it showed that subjects generally agree about the most appropriate moments to tap in time when listening to a passage of music.
10. Can anyone describe to me what they were thinking while they were ‘deciding’ when to tap? (Assuming not...) This shows that most of you are not consciously aware of the mental processes involved when you tap in time to a passage of music.

4. 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

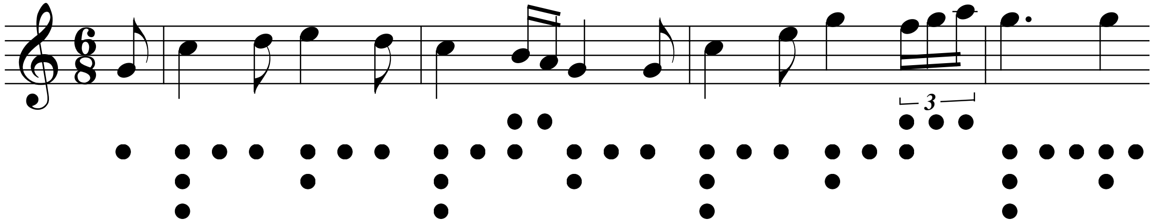
4. Lerdahl and Jackendoff's (1983) *Generative Theory of Tonal Music*

1. For the remainder of this lecture I'm going to introduce you to a number of different theories of metre perception that have been developed over the past twenty years or so.
2. I'm going to start with what is probably the most famous theory of music perception and cognition, namely, Lerdahl and Jackendoff's (1983) *Generative Theory of Tonal Music.* From this point on I'll refer to this theory as *GTTM.*
3. *GTTM* is essentially a system that takes a representation of a passage of music as input and generates as output a structural description of the passage that is supposed to correctly describe certain aspects of how the passage is interpreted by an expert listener.
4. Lerdahl and Jackendoff call the input representation to their theory the 'musical surface'. One of the problems with their theory is that it's not at all clear what they intend this musical surface to be: for example, in some places it seems that they're assuming it's the sound heard by the listener and in others it seems they're assuming that the musical surface is a score.
5. *GTTM* consists of four interacting components:
    (a) The first component generates a description of what Lerdahl and Jackendoff call the 'grouping structure' of a passage. This is the way that a passage is heard to be split up into motives, themes, phrases, sections and so on.
    (b) The second component consists of a set of rules that work together to generate a metrical structure for a passage. Lerdahl and Jackendoff (1983, p. 17) define the 'metrical structure' of a passage to be 'the regular, hierarchical pattern of beats to which the listener relates musical events'.
    (c) The third component of the theory generates what Lerdahl and Jackendoff call a 'time-span reduction' of a passage. The time-span reduction component uses the grouping and metrical structure of a passage to predict the perceived relative importance of events within a passage—i.e., which events are perceived to be elaborations of or dependent on or subordinate to other events.

   (d) The fourth component of Lerdahl and Jackendoff's (1983) 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 primarily by the harmonic movement within the passage.

6. 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 those structural descriptions in the set defined by the well-formedness rules that best describe how an expert listener interprets the passage given to the theory as input.

5. Lerdahl and Jackendoff's (1983) Metrical Well-Formedness Rules



MWFR 1 Every attack point must be associated with a beat at the smallest metrical level present at that point in the piece.

MWFR 2 Every beat at a given level must also be a beat at all smaller levels present at that point in the piece.

MWFR 3 At each metrical level, strong beats are spaced either two or three beats apart.

MWFR 4 The tactus and immediately larger metrical levels must consist of beats equally spaced throughout the piece. At subtactus metrical levels, weak beats must be equally spaced between the surrounding strong beats.

5. Lerdahl and Jackendoff's (1983) Metrical Well-Formedness Rules

1. Let's take a closer look at the metrical structure component of Lerdahl and Jackendoff's (1983) *GTTM*, beginning with the metrical well-formedness rules of which there are four.

2. The first of these rules states that "every attack point must be associated with a beat at the smallest metrical level present at that point in the piece" (Lerdahl and Jackendoff, 1983, pp. 72, 347).

3. This simply expresses the idea that every timepoint at which an event occurs in a piece must have *some* metric strength even if it is a very weak beat.

4. The second rule states that "every beat at a given level must also be a beat at all smaller levels present at that point in the piece" (Lerdahl and Jackendoff, 1983, pp. 72, 347).

5. This rule expresses one of the things that we discovered in our little experiment, namely that if you tap in time to a piece of music at a particular speed, you do not miss out any of the timepoints that you would tap at if you had to tap at a slower speed.

6. So you can see here in this example that all the tactus beats are also beats at the subtactus level and all the supertactus beats (which occur at the barlines) are also beats at the tactus level.

7. The third metrical well-formedness rule states that "at each metrical level, strong beats are spaced either two or three beats apart" (Lerdahl and Jackendoff, 1983, pp. 69, 347).

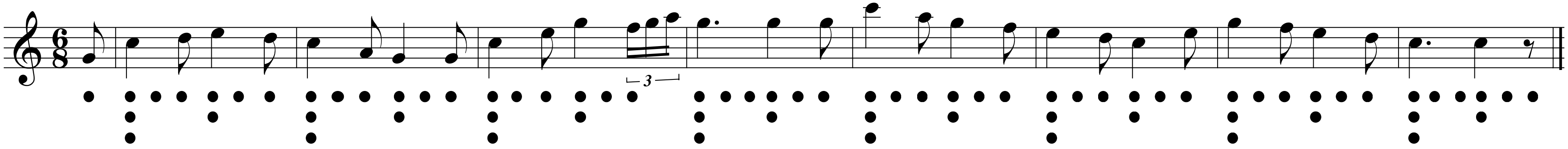
8. So, again, in this example, you can see that the tactus level beats are spaced three subtactus level beats apart and the supertactus level beats are spaced two tactus-level beats apart.

9. Lerdahl and Jackendoff acknowledge that this rule only applies to Western tonal classical music (and not even to all of that—e.g., third movement of Brahms C minor trio, Op. 101 is in seven and the second movement of Tchaikovsky's *Pathétique* Symphony is in five).

10. The fourth rule is also specific to Western tonal classical music and states that “the tactus and immediately larger metrical levels must consist of beats equally spaced throughout the piece. At subtactus metrical levels, weak beats must be equally spaced between the surrounding strong beats” (Lerdahl and Jackendoff, 1983, pp. 72, 347).

11. In their first version of this rule, Lerdahl and Jackendoff specified that *every* metrical level must be evenly spaced. But this would fail to account for the very commonly occurring phenomenon of triplets. For example, two levels below the tactus in this melody we have a quaver divided into two semiquaver beats here and a quaver divided into three triplet semiquaver beats here. So the semiquaver metrical level in this melody does not consist of beats that are equally spaced thoughout the passage. However, both here and here the weak beats are equally spaced between the stronger beats as required by the fourth metrical well-formedness rule.

6. Lerdahl and Jackendoff's (1983) Metrical Preference Rules 1–4



MPR 1 (Parallelism) Where two or more groups or parts of groups can be construed as parallel, they preferably receive parallel metrical structure.

MPR 2 (Strong Beat Early) Weakly prefer a metrical structure in which the strongest beat in a group appears relatively early in the group.

MPR 3 (Event) Prefer a metrical structure in which beats of level L_i that coincide with the inception of pitch-events are strong beats of L_i.

MPR 4 (Stress) Prefer a metrical structure in which beats of level L_i that are stressed are strong beats of L_i.

6. Lerdahl and Jackendoff's (1983) Metrical Preference Rules 1–4

1. The metrical well-formedness rules simply tell us which metrical structures are possible.
2. Let's go on now and look at the Metrical Preference Rules which predict which of the possible metrical structures best describe our perception of a given passage.
3. The first metrical preference rule is called the 'Parallelism' rule and it predicts that "where two or more groups or parts of groups can be construed as parallel, they preferably receive parallel metrical structure" (Lerdahl and Jackendoff, 1983, pp. 75, 347).
4. For example, in this melody here, the first two-bar phrase and the second two-bar phrase both begin with a rising fourth, followed by a rising contour and both begin with the same rhythm. This makes us perceive them as 'parallel'—that is, similar in some significant way and the first metrical preference rule predicts that this fact that we perceive them as parallel makes us want to assign them parallel metrical structures so that, for example, this C_5 will have a similar metrical strength to this C_5 and this E_5 will have a similar metrical strength to this D_5 and so on.
5. The second metrical preference rule is called the 'Strong Beat Early' rule and it predicts that we "weakly prefer a metrical structure in which the strongest beat in a group appears relatively early in the group" (Lerdahl and Jackendoff, 1983, pp. 76, 347).
6. This is obviously related to the fact that the vast majority of pieces of tonal music begin either at the beginning of a bar or with a very short anacrusis.
7. Note that both of these first two metrical preference rules focus on the effect of grouping on metrical structure, grouping being the way that a passage is perceived to be split up into phrases, motives, sections and so on. I'll talk more about grouping later on in this course.
8. The third metrical preference rule is called the 'Event' rule and it predicts that we should "prefer a metrical structure in which beats of level L_i that coincide with the inception of pitch-events are strong beats of L_i" (Lerdahl and Jackendoff, 1983, pp. 76, 347).

9. Basically, this rule just predicts that we prefer not to hear a rhythm as being syncopated if we can get away with it. So, here, for example, we hear the tactus beats as falling at the attack points of the crotchets and not on the second and fourth quavers of each bar where nothing happens. If we heard the second and fourth quavers of each bar as being strong beats then almost every event in the melody would be heard as being a syncopation.

10. Of course, syncopations are not infrequent in tonal music but the third metrical preference rule predicts that we perceive a passage of tonal music in a way that minimises the number of syncopations. Temperley (2001) has actually shown that this is not the case for African and rock music.

11. The fourth metrical preference rule is called the 'Stress' rule and it predicts that we 'prefer a metrical structure in which beats of level L_i that are stressed are strong beats of L_i" (Lerdahl and Jackendoff, 1983, pp. 79, 347).

12. This is a fairly self-evident rule that simply predicts that relatively louder events will be heard as being relatively metrically stronger. In fact, it seems that the loudness of events has surprisingly little effect on the perception of metrical structure. Most of you, for example, were perfectly capable of detecting the metric structure of this melody even when all the notes had exactly the same timbre and loudness.

7. Lerdahl and Jackendoff's (1983) Metrical Preference Rule 5

MPR5 (Length) Prefer a metrical structure in which a relatively strong beat occurs at the inception of either

a. a relatively long pitch-event,

b. a relatively long duration of a dynamic,

c. a relatively long slur,

d. a relatively long pattern of articulation,

e. a relatively long duration of a pitch in the relevant levels of the time-span reduction, or

f. a relatively long duration of a harmony in the relevant levels of the time-span reduction (harmonic rhythm).

7. Lerdahl and Jackendoff's (1983) Metrical Preference Rule 5

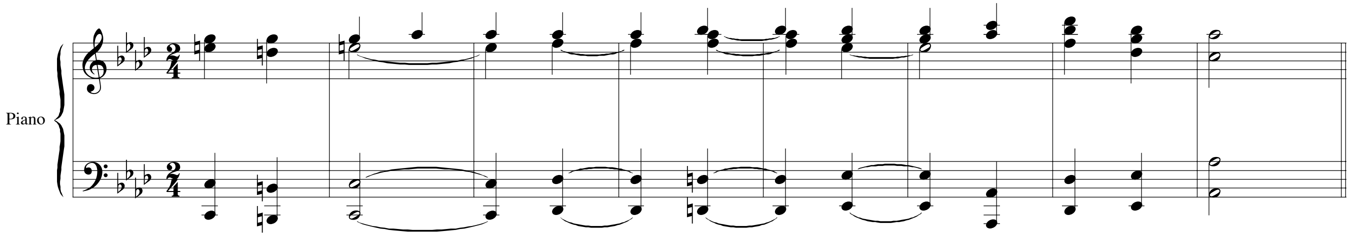
1. The fifth metrical preference rule predicts that we "prefer a metrical structure in which a relatively strong beat occurs at the inception of either
   - a. a relatively long pitch-event,
   - b. a relatively long duration of a dynamic,
   - c. a relatively long slur,
   - d. a relatively long pattern of articulation,
   - e. a relatively long duration of a pitch in the relevant levels of the time-span reduction, or
   - f. a relatively long duration of a harmony in the relevant levels of the time-span reduction (harmonic rhythm).
2. To test MPR5a, I want you to listen to two sequences of notes and make note of which tones in the sequence you hear as being on strong beats.
   - (a) [PLAY MPR5a1] MPR4 would predict that you heard the loud note as occurring on a downbeat. How many of you heard it that way?
   - (b) [PLAY MPR5a2] Who heard the loud note as occurring on the strong beat? Who heard the long note as occurring on the strong beat? That sequence was derived from the first simply by deleting two of the quieter notes and lengthening the first quiet one to fill the gap. According to Lerdahl and Jackendoff's theory, there should now be a conflict between MPR4 and MPR5: MPR4 predicts that you should hear the strong beat on the loud note but MPR5a predicts that you should hear the strong beat on the long note.
3. MPR5b would predict that the strong beat in the first of the sequences that you've just heard should be heard to occur on the first quiet note—which doesn't seem to be in agreement with what most of you perceived.
4. To test MPR5c and MPR5d, I want you now to listen to another sequence and make note of which notes in the sequence occur on the strong beats.

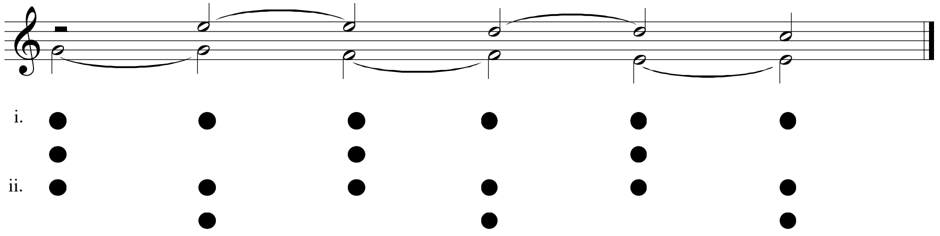
   (a) [PLAY MPRd] How many of you heard the strong beat as falling on the staccato note? How many of you heard the strong beat as falling on the first legato note?

   (b) MPR 5d predicts that the strong beat in this example should fall on the first legato note in each sequence of three legato notes.

5. MPR 5e and 5f relate to time-span reduction which I'm not going to consider today so we'll move on to the rest of the Metrical Preference Rules.

8. Lerdahl and Jackendoff's (1983) Metrical Preference Rules 6–10





MPR 6 (Bass) Prefer a metrically stable bass.

MPR 7 (Cadence) Strongly prefer a metrical structure in which cadences are metrically stable; that is, strongly avoid violations of local preference rules within cadences.

MPR 8 (Suspension) Strongly prefer a metrical structure in which a suspension is on a stronger beat than its resolution.

MPR 9 (Time-Span Interaction) Prefer a metrical analysis that minimizes conflict in the time-span reduction.

MPR 10 (Binary Regularity) Prefer metrical structures in which at each level every other beat is strong.

8. Lerdahl and Jackendoff's (1983) Metrical Preference Rules 6–10

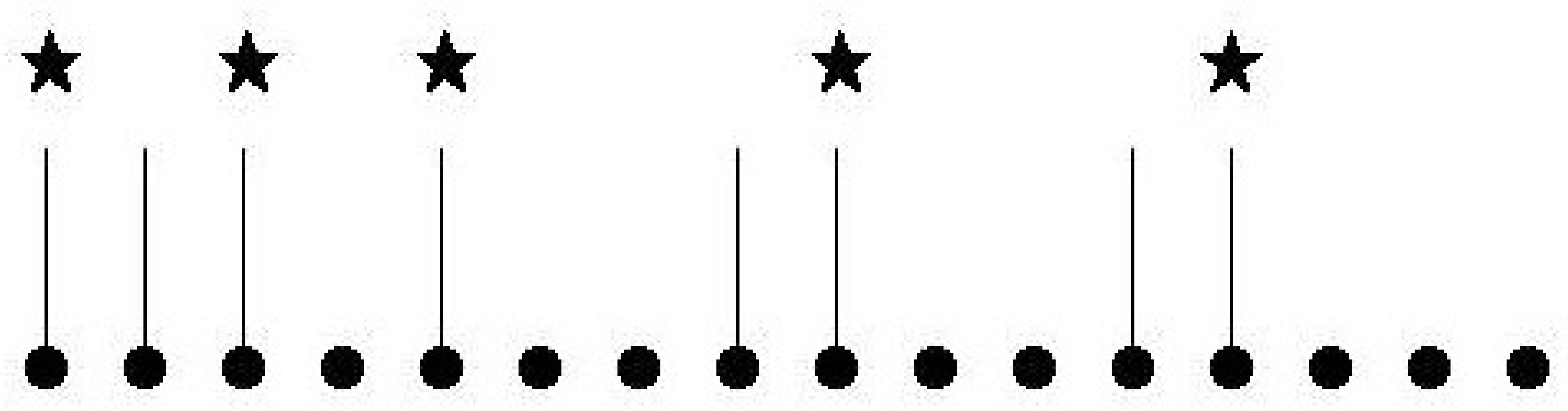
1. MPR6 predicts that we "prefer a metrically stable bass" (Lerdahl and Jackendoff, 1983, pp. 88, 348).
2. The purpose of this rule is to intensify the effect of MPRs 3 (the Event rule), 4 (the Stress rule) and 5 (the Length rule) when these rules apply to the bass line.
3. MPR7 states that we "strongly prefer a metrical structure in which cadences are metrically stable; that is, [we] strongly avoid violations of local preference rules within cadences" (Lerdahl and Jackendoff, 1983, pp. 88, 348).
4. Lerdahl and Jackendoff admit that MPR7 is just a 'preliminary statement' as it leaves a lot unsaid: for example, they do not specify how we are supposed to identify cadences in the musical surface, nor do they specify how we are supposed to determine whether a cadence is a 'masculine ending' implying a weak-strong accent pattern or a feminine-ending implying a strong-weak accent pattern.
5. In support of their MPR 7, Lerdahl and Jackendoff give the following extract from the second movement of Beethoven's Sonata Op.110 in which the cadence is almost the only point of metrical stability in the phrase. [PLAY Beethoven OP110].
6. MPR8 proposes that we "strongly prefer a metrical structure in which a suspension is on a stronger beat than its resolution" (Lerdahl and Jackendoff, 1983, pp. 89, 348).
7. To understand this rule, we're going to do another little experiment. First listen to this short fragment a couple of times [Play suspension4.38a twice]. Now I'm going to play two versions of the fragment, each with a drum beat to mark the metrically strong beats. I want you to decide which of these two versions you think sounds more natural. [Play suspension4.38ai then suspension4.38aii twice each]. Who thought the first version was the more natural of the two? Who thought the second version sounded more natural?
8. MPR8 would predict that the first version would sound more natural. Also MPR6 would support that interpretation by predicting that the first note which is in the lower part should be heard as being metrically stable.
9. The ninth metrical preference rule concerns the interaction between metrical structure and the time-span reduction component of GTTM which I'm not going to discuss right now.

10. Finally, the tenth metrical preference rule—the so-called "Binary Regularity" rule—predicts that we "prefer metrical structures in which at each level every other beat is strong" (Lerdahl and Jackendoff, 1983, pp. 101, 348).
11. This rule sounds like it predicts that we are generally biased in favour of duple metres rather than triple metres, which may be the case. However, it was actually motivated by Lerdahl and Jackendoff's desire to account for 'hypermetre'—that is, metrical regularities at levels higher than the bar level. They believe that some bar-line timepoints are metrically stronger than others. However, for this to be the case, the absolute duration of each bar must be less than about 1800ms (Parncutt, 1994).
12. To sum up, Lerdahl and Jackendoff provided the first reasonably complete and at least semi-formalized theory of metrical structure perception in Western tonal music.
13. They adopted an architecture for their theory in which the rules were expressed as preference rules that are supposed to work together to select an optimal description of the metrical structure of a passage that is perceived by a listener.
14. Main problem with applying the preference rules is that Lerdahl and Jackendoff don't propose any way of weighting the rules in order to resolve rule conflicts.
15. Another major problem with GTTM is the circular dependence of one component of the theory on another. For example, the rules in the grouping-structure component use information in the time-span reduction but the rules in the time-span reduction component use information from the grouping structure analysis. Therefore actually implementing the theory as a computer program would be a very major undertaking. There have been a number of attempts to implement parts of the theory, however. For example, David Temperley's computational model of metre perception is, essentially, an implementation of a significant chunk of Lerdahl and Jackendoff's metrical structure theory.
16. After the break I'll talk about some of the theoretical and experimental work that's been done on metre perception since the publication of GTTM.

9. Povel and Essens's (1985) theory of temporal pattern perception

A tone in a sequence will sound accented if:

1. it is relatively isolated,
2. it is the second tone in a cluster of two tones, or
3. it is the first or last tone in a cluster of 3 or more tones.

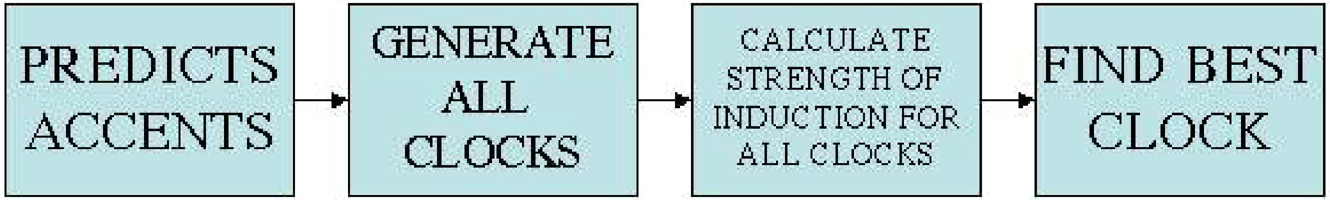


9. Povel and Essens's (1985) theory of temporal pattern perception

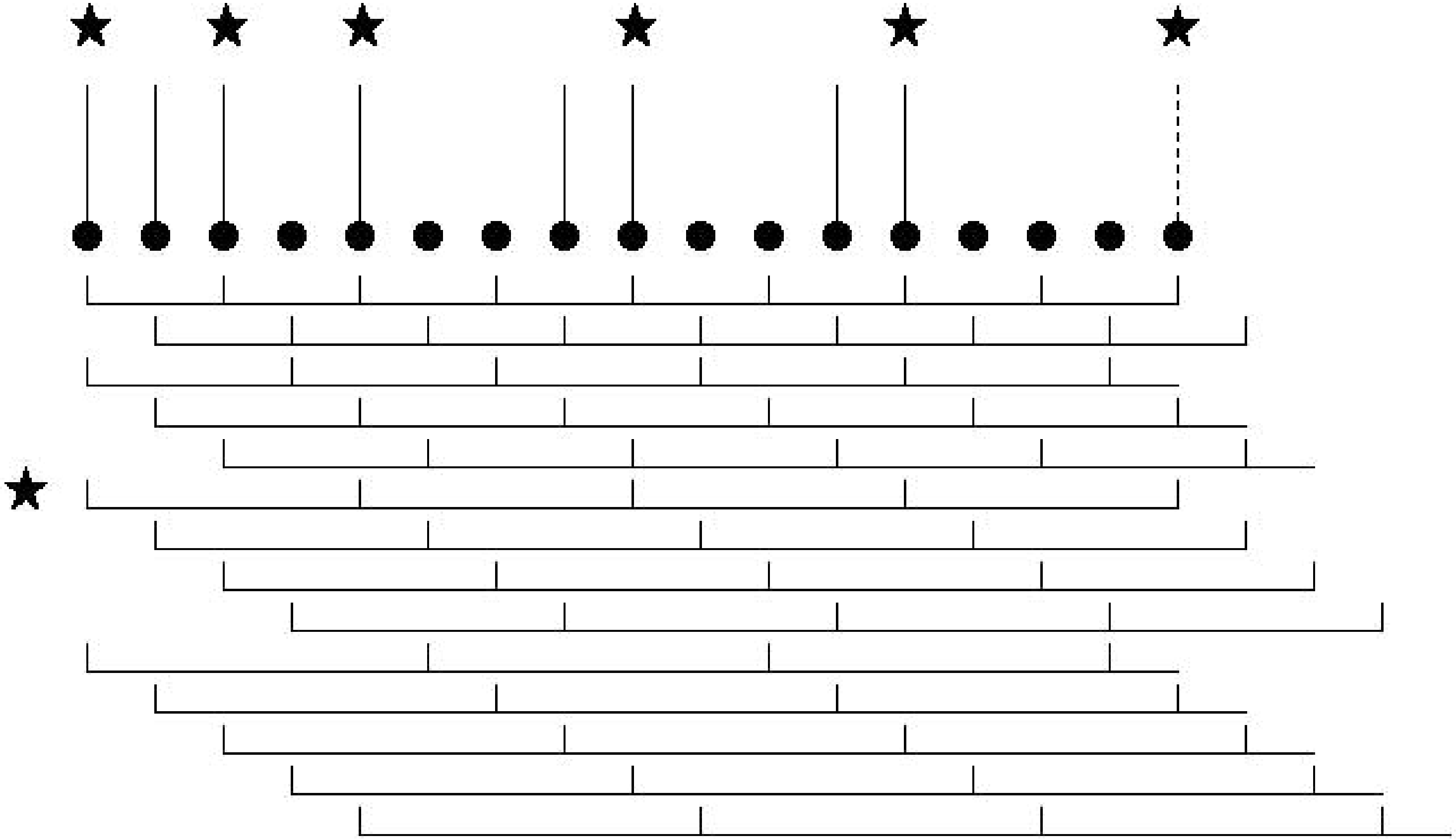
1. We've seen that most listeners find it very easy and natural to tap or clap a regular beat in time to many passages of music. Moreover, most listeners seem to agree on the most appropriate places at which to clap or tap and we call this set of most natural places to tap, the tactus.
2. In a famous paper published in 1985, Povel and Essens explain this behaviour by hypothesizing that when a listener hears a repeated sequence of events forming a repeated rhythmic pattern, this induces an 'internal clock' to start 'ticking' in their minds at a particular rate, such that the ticks occur at what we've been calling the tactus timepoints.
3. Povel and Essens studied this 'clock-induction' behaviour by playing subjects repeated patterns of tones in which all the tones had the same pitch, loudness, duration and timbre and differed only in their onset times. Here's an example of one of the sequences that Povel and Essens used in their experiments. [Play PECAT1]
4. Povel and Essens observed that listeners have a strong impression that some of the notes are accented even in sequences where all the tones have the same loudness, pitch and timbre, like the one that I just played you. On your handout, you'll see that I've represented the sequence as a sequence of vertical strokes separated by dots which represent empty beats. As you listen to the sequence again, I want you to mark on the handout the tones that sound most accented [Play PECAT1 again].
5. [Put on SLIDE 9 with bottom covered up]
6. Povel and Essens devised three simple rules for predicting which tones will sound accented in a sequence of identical tones like the one I just played you. They predicted that a tone in such a sequence will sound accented if
   (a) it is relatively isolated,
   (b) it is the second tone in a cluster of two tones,
   (c) it is the first or last tone in a cluster of three or more tones.
7. [Uncover bottom half of SLIDE 9]

8. For example, these rules predict that in the sequence I just played you, the accented tones will be the ones marked here in this diagram.
9. Who heard precisely these tones as being accented? Who heard some of these tones as being accented and no others? Who heard different tones as being accented?
10. Povel and Essens hypothesised that it is the distribution of accented tones in such a sequence that determines the internal clock (or tactus) induced in the listener and they devised a model, implemented as a computer program, that predicts the tactus that a listener will tap when they hear a repeated rhythmic sequence like the one I just played you.

10. Povel and Essens's (1985) Model



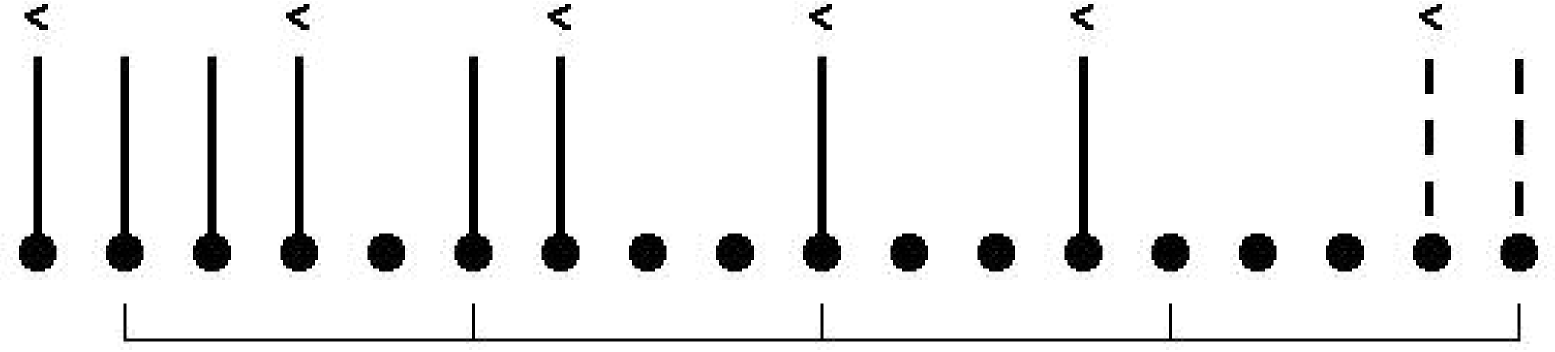
PREDICTS ACCENTS → GENERATE ALL CLOCKS → CALCULATE STRENGTH OF INDUCTION FOR ALL CLOCKS → FIND BEST CLOCK



10. Povel and Essens's (1985) Model

1. The first step in this model is to predict which tones in the sequence will be heard as being accented using these three rules.
2. Next, the program generates all the possible clocks with a period less than half the length of the repeated rhythmic pattern. In all the sequences used by Povel and Essens, the onset time of every tone is an integer multiple of 200ms after the beginning of the sequence. Therefore, Povel and Essens only consider clocks with ticks that occur an integer multiple of 200ms after the beginning of the sequence. For example, this diagram shows all the possible clocks for the sequence that I played you a couple of minutes ago.
3. The next step in the model is to calculate for each possible clock a quantity that is supposed to measure how strongly that particular clock is suggested or induced by the rhythmic sequence.
4. Povel and Essens propose that for a clock to be strongly induced, many of the clock ticks must coincide with accented notes in the sequence. The more clock ticks that coincide with gaps or unaccented notes, the less strongly that clock will be induced. The clock that their model predicts will be most strongly induced by a rhythmic sequence is then the one with the smallest number of ticks coinciding with gaps and unaccented notes.
5. The clock that Povel and Essens' theory predicts will be most strongly induced by the tone sequence I just played you is shown here in this diagram and it's indicated by the drum beat in this recording. [Play PECAT1_WITH_DRUM].

A pattern that does not induce any clock strongly



- *First experiment*:
  - Showed that subjects find it easier to learn rhythmic sequences that strongly induce internal clocks.
  - The intervals between consecutive events in sequences that do not induce a clock are *figurally coded*, that is, each interval is categorized as either “long” or “short”.
- *Second experiment*:
  - Showed that subjects find it easier to learn rhythmic sequences when they are accompanied by a regular bass tone indicating the most strongly induced internal clock.

A pattern that does not induce any clock strongly

1. Now it turns out that for some rhythmic sequences, even the most strongly induced clock is still only very weakly induced because there is no clock in which many ticks coincide with accented tones. Here's an example of a sequence that does not strongly induce any clock. [Play PECAT7]

2. For this sequence, Povel and Essens' theory predicts that the clock shown here will be the most strongly induced. Here's what this sequence sounds like with this clock indicated by drum beats. [Play PECAT7_WITH_CLOCK].

3. Povel and Essens carried out three experiments to test their model.

4. In the first experiment, they tested the hypothesis that the ease with which a subject learns a rhythmic pattern and the accuracy with which a subject can reproduce the pattern are greater for rhythmic patterns that more strongly induce an internal clock.

5. In this experiment, subjects could listen to a repeated rhythmic pattern as many times as they liked and then they had to reproduce it by tapping it out. PE noted how many times the subject had to listen to the pattern before he or she felt they were ready to attempt to reproduce it. They also measured how accurately each subject reproduced the pattern.

6. PE used their model to predict how strongly each of the tone sequences used in the experiment induced an internal clock and they found that there was a significant correlation between how easily subjects learned a pattern and the strength with which it induced an internal clock, as calculated by their model.

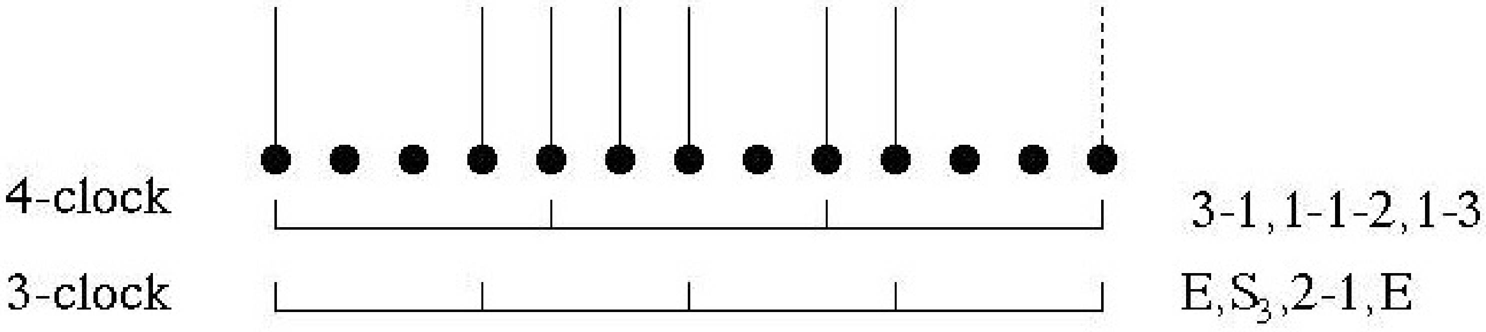
7. They found that the patterns that were most difficult to learn seemed not to induce *any* clock whatsoever. Subjects seemed to remember these sequences by dividing up the sequence into groups of tones and categorising the intervals between tones into just two approximate categories, "long" and "short". Povel and Essens call this 'figural coding' as opposed to the 'durational coding' that is assumed to occur when a subject remembers a sequence by inducing an internal clock.

8. In their second experiment, subjects again had to learn and reproduce rhythmic sequences but in this experiment, each tone sequence was presented accompanied by a low bass tone on each tick of the most strongly induced

clock for that sequence. Here's an example of the kind of stimulus used in this second experiment. [Play PILOT_EXP_3_4_clock].

9. Povel and Essens found that subjects were much better at remembering the rhythms when they were accompanied by a pulse indicating the tactus or internal clock in this way.
10. The fact that listeners find it easier to recall rhythms when they are presented with a 'clock-like' accompanying beat suggests that we use the internal clock to encode the rhythm.
11. Before describing Povel and Essens' third experiment, I'm going to play you two tone sequences. In each tone sequence, you'll hear a rhythmic pattern played using a high-pitched tone accompanied by a clock-like, equally-spaced bass tone. Listen carefully to the two sequences. [PLAY PILOT_EXP_3_4_clock and PILOT_EXP_3_3_clock] Put your hand up if you think that the top part in each of those examples was the same. Put your hand up if you think that the top parts in those two sequences were different.
12. Povel and Essens carried out this experiment and found that 9 out of 10 people were not able to recognize that the top parts in 'double-sequences' like that were the same when the bass part indicated a different metre or internal clock.
13. This suggests that the way that the rhythmic sequence is encoded in the listener's mind depends upon the internal clock that is induced: if you induce a different clock under the same sequence then the sequence will be encoded differently and listeners will not recognize that the same sequence has been presented.

11. Povel and Essens's (1985) encoding system



1. If a clock unit is subdivided into equal intervals, this is described in the code by means of the symbol *S* and a subscript indicating the number of intervals in that unit.
2. If a clock unit is empty, this is indicated with the symbol *E*.
3. If a clock unit is subdivided in intervals of unequal length, no reduction is possible and the successive intervals are described by indicating which proportion of the clock unit they take.

11. Povel and Essens's (1985) encoding system

1. Povel and Essens propose a tentative encoding language in which a rhythm is encoded using the following three rules:
   (a) If a clock unit is subdivided into equal intervals, this is described in the code by means of the symbol S and a subscript indicating the number of intervals in that unit.
   (b) If a clock unit is empty, this is indicated with the symbol E.
   (c) If a clock unit is subdivided in intervals of unequal length, no reduction is possible and the successive intervals are described by indicating which proportion of the clock unit they take.
2. This diagram here shows how a single rhythmic pattern will be encoded in terms of different clocks using these encoding rules. Note that the encoding of the sequence with a clock of period 3 is shorter and therefore more efficient than the encoding that results with a clock of period of 4.
3. Povel and Essens hypothesized that in this sort of situation, a listener will judge that a sequence sounds simpler when it is accompanied by the clock that results in the shorter and more efficient encoding. For example, in the example I played you just now, the one where the melody is accompanied by a clock of period 3 should sound simpler. Listen once again to these two sequences and see if you agree with that. [PLAY PILOT_EXP_3_4_clock and PILOT_EXP_3_3_clock]
4. How many thought the first sounded simpler? How many thought the second sounded simpler?
5. In their third experiment, Povel and Essens played subjects pairs of stimuli like that and got the subjects to judge which stimulus in the pair sounded simpler. They found that the results supported their hypothesis that listeners find that a sequence sounds simpler when it is accompanied by a clock that allows the sequence to be encoded more efficiently using their encoding system.

12. Lee's (1991) theory of metre perception

- Models of Lee (up to 1985), Longuet-Higgins and Steedman only generate *one* metrical structure for any given passage.
- Lee identifies 4 basic assumptions in earlier work:
  1. Longer notes begin stronger beats.
  2. Prefer not to hear a melody as being syncopated (Lerdahl and Jackendoff, 1983; Longuet-Higgins and Lee, 1984; Povel and Essens, 1985).
  3. Repetition gives us important cues about the metrical structure.
  4. Theories of Longuet-Higgins, Lee and Steedman assume that the fact that listeners build up their interpretation of the metrical structure *as they listen to the music* strongly affects how they perceive the metre.

12. Lee's (1991) theory of metre perception

1. In his PhD thesis which he completed in 1989, Christopher Lee does a number of experiments to test some of the basic assumptions underlying several earlier models of metre perception and presents a new model that incorporates the findings of these experiments.
2. This dissertation is summarised in Lee 1991 where he begins by presenting a critical review of a number of earlier theories of metre perception including those of Lerdahl and Jackendoff and Povel and Essens that I've already described together with a number of earlier models developed by himself, Longuet-Higgins and Steedman.
3. One important difference between the models developed by Longuet-Higgins, Steedman and Lee up to 1985 and those of Povel and Essens and Lerdahl and Jackendoff is that the models of Longuet-Higgins, Steedman and Lee only ever predict a single most likely metrical structure for any given input passage.
4. Lee identifies a number of basic assumptions in these theories:
    - (a) Nearly all the theories assume that longer notes are more likely to be perceived as beginning on metrically stronger beats.
    - (b) Some of the theories (Lerdahl and Jackendoff, 1983; Longuet-Higgins and Lee, 1984; Povel and Essens, 1985) assume that we prefer to hear strong beats as coinciding with event onsets—that is, that we prefer not to interpret a melody as being syncopated unless we have to.
    - (c) All the theories assume that the occurrence of repeated note patterns gives us important clues as to the metrical structure that we are intended to perceive.
    - (d) The theories of Longuet-Higgins, Steedman and Lee, unlike those of Povel and Essens and Lerdahl and Jackendoff (1983), assume that "the listener's final choice of interpretation for a given sequence depends on the history of his or her metrical judgements during the course of listening to the sequence" (Lee, 1991, p. 63). In other words, these theories assume that the metrical structures that we induce during the listening process depend upon the fact that the evidence that we use to induce these structures is only gradually revealed to us as the music unfolds.

13. Summary of early theories of Longuet-Higgins, Lee and Steedman

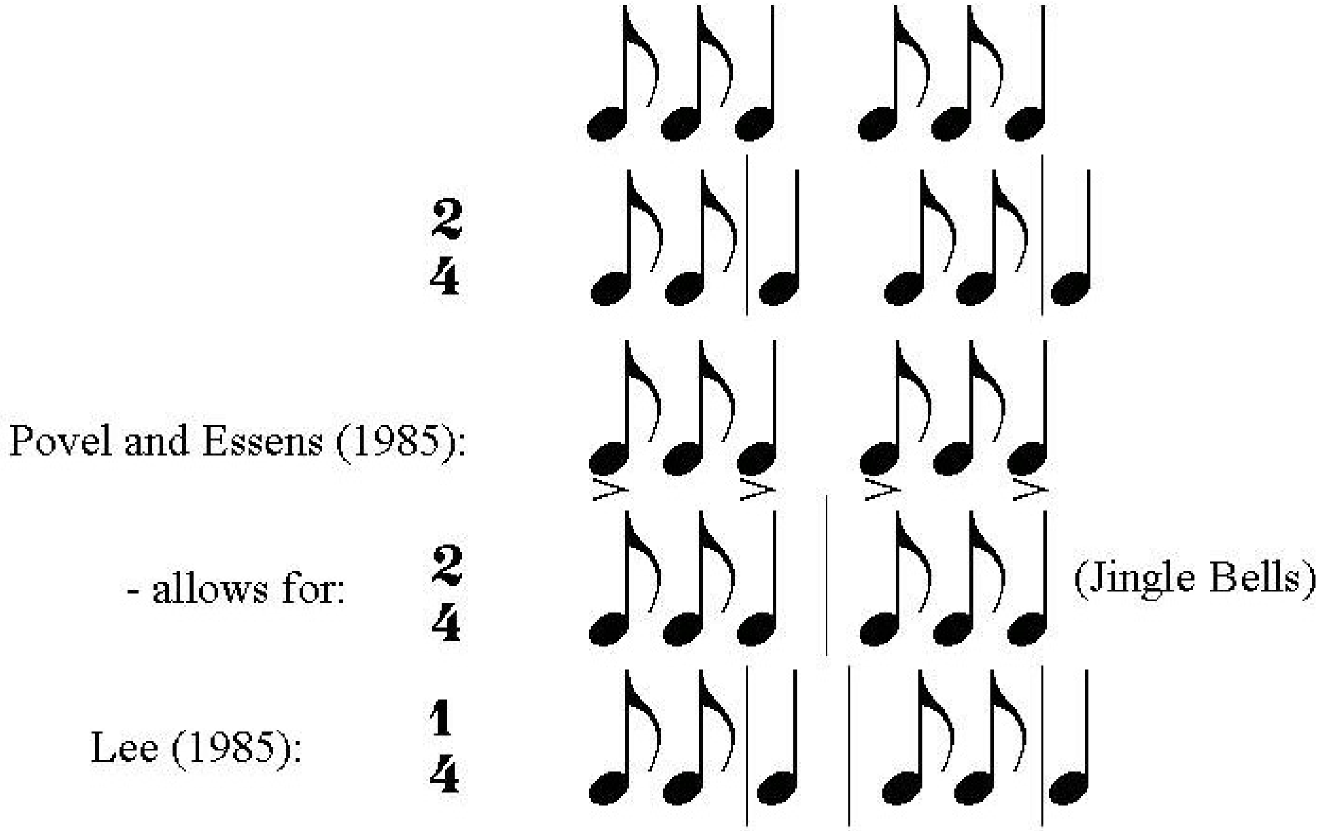
- Longuet-Higgins and Steedman (1971)
  - assumes listeners build up metrical interpretation as they listen
  - based on observation that dactyl rhythm nearly always starts on a strong beat
- Steedman (1973, 1977)
  - extension of Longuet-Higgins and Steedman (1971)
  - searches for repeated melodic figures and assumes that period between beginnings of consecutive occurrences is a metrical unit
- Longuet-Higgins and Lee (1982)
  - another extension of the algorithm that assumes that longer notes begin on stronger beats
- Longuet-Higgins and Lee (1984)
  - development of Longuet-Higgins and Lee (1982)
  - incorporates assumption that listeners try to find an unsyncopated interpretation of a passage
- Lee (1985)
  - development of Longuet-Higgins and Lee (1982)
  - can infer more than one metrical level
  - disallowed from inferring metrical units longer than longest note in melody

13. Summary of early theories of Longuet-Higgins, Lee and Steedman

See slide.

14. Long notes and metrical structure

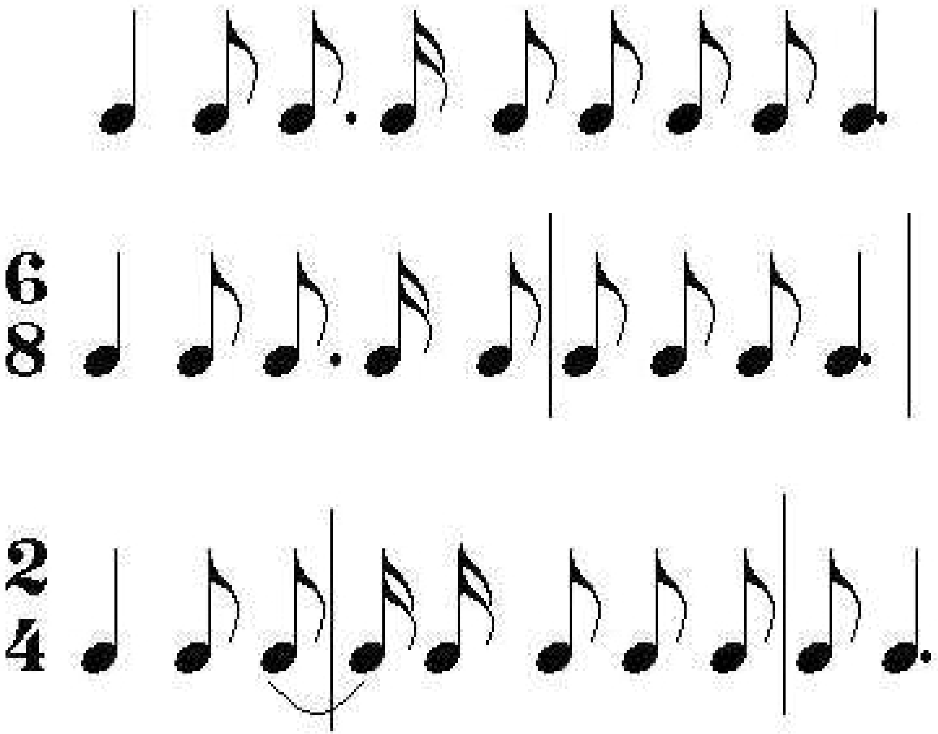
Longuet-Higgins and Steedman (1971), Steedman (1973, 1977), Longuet-Higgins and Lee (1982): all based heavily on assumption that longer notes are heard as beginning on stronger beats.



14. Long notes and metrical structure

1. The models of Longuet-Higgins and Steedman (1971), Steedman (1973, 1977), Longuet-Higgins and Lee (1982) are all heavily based on the idea that listeners hear longer notes as starting on stronger beats.
2. All these models predict, for example, that we interpret this rhythm [Play Jingle Bells] as having the metric structure shown here.
3. Lerdahl and Jackendoff's (1983) fifth metrical preference rule—the Length rule—would also predict that we hear the rhythm in this way.
4. However, Povel and Essens's (1985) accent based theory of clock induction predicts that we would hear this rhythm in the way shown here and this seems a very reasonable prediction—indeed, this is precisely the way it is notated at the beginning of Jingle Bells.
5. So it seems that it might be dangerous to place too much weight on the idea that longer notes are generally heard to begin on stronger beats.
6. Lee (1985) tries to control the long-note-strong-beat effect by not allowing his model to infer 'internal clocks' or metrical levels in which the periods between ticks are longer than the longest note in the music. So in the Jingle Bells example, the highest metrical level inferred by Lee's (1985) model would be the crotchet level—that is, it would not predict any particular placement for the barlines.

15. Avoiding syncopation

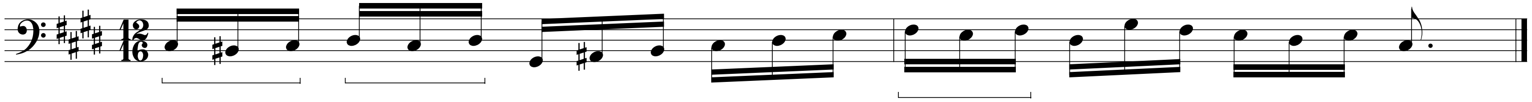


- This implies that we strongly avoid syncopations (Longuet-Higgins and Lee, 1984).
- Syncopation = "occurrence, on a rest or tied note of a beat that is stronger than the one on the immediately preceding sounded note" (Lee, 1991, p. 70).

15. Avoiding syncopation

1. Recall that Lerdahl and Jackendoff's (1983) third Metrical Preference Rule, the Event rule, predicted that listeners select a metrical structure that maximises the number of strong beats that coincide with pitch-events.
2. As Lee (1991) points out, however, an unintended implication of this rule is that we prefer duple rhythms to triple rhythms because strong beats occur more frequently in a duple metre than in a triple metre.
3. For example, Lerdahl and Jackendoff's (1983) theory would predict that this rhythm is just as likely to be heard as being in 2/4 as in 6/8. Listen to it and see which of the two interpretations you think is more natural. [PLAY LEE30]
4. Longuet-Higgins and Lee (1984) avoid this problem by building into their model a rule that causes it to strongly avoid what they call 'major syncopations'. Lee defines a syncopation as being "the occurrence, on a rest or tied note, of a beat that is stronger than the one on the immediately preceding sounded note" (Lee, 1991, p. 70).

16. Repetition and metrical structure

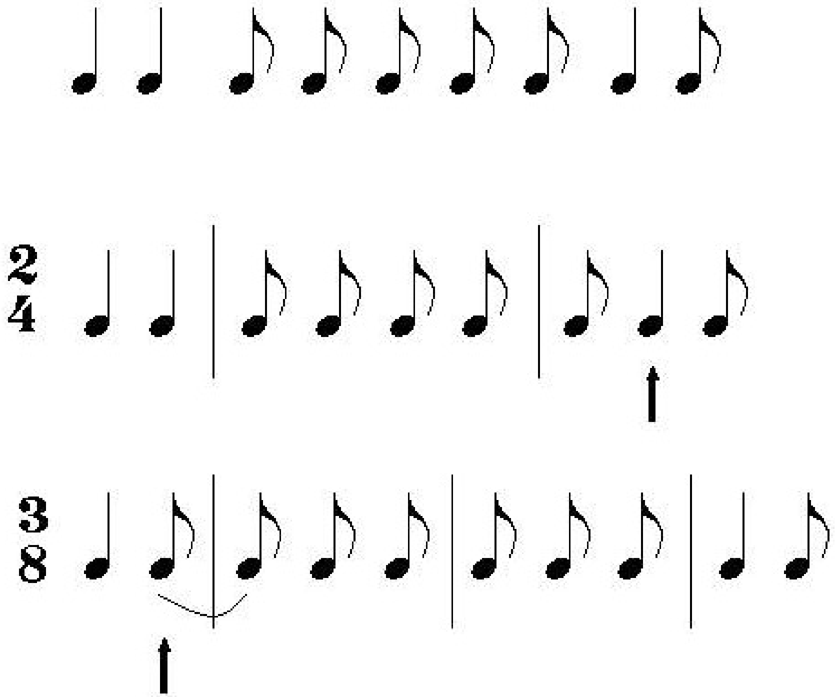


Lerdahl and Jackendoff (1983) and Steedman (1973, 1977) incorporate idea that repetition strongly affects the metrical structure that is induced.

16. Repetition and metrical structure

1. All the theories and models reviewed by Lee acknowledge that repetition of melodic patterns strongly influences how we hear the metrical structure of a melody.
2. However, it is only the theories of Steedman (1973, 1977) and Lerdahl and Jackendoff (1983) that actually attempt to incorporate this assumption into their models.
3. Steedman proposes that if we hear a melodic pattern as being a repeat of a previous pattern then we take the period between the beginnings of the patterns to be a metrical unit.
4. For example, Steedman's model predicts that in this passage taken from the beginning of the fourth fugue from Book 2 of Bach's 48, the first repeat of this lower-neighbour note pattern induces the dotted-quaver metrical level and the second repeat here induces the bar-level. [Play WTCII4b]

17. Processing considerations



- Lerdahl and Jackendoff (1983), Longuet-Higgins and Lee (1984), Povel and Essens (1985) assume that listener does not infer metrical structure until he/she has heard the complete passage.
- Longuet-Higgins and Steedman (1971), Steedman (1973, 1977) assume that listeners listen hard at the beginning of a passage to induce metrical structure and then need strong counter-evidence to make them change their minds.

17. Processing considerations

1. Lee points out that a major problem with the theories of Lerdahl and Jackendoff (1983), Longuet-Higgins and Lee (1984) and Povel and Essens (1985) is that they effectively assume that a listener does not infer any metrical structure until they have heard a complete passage: that is, they do not take into account the fact that a listener builds up his or her interpretation of the metrical structure as the passage unfolds in time.
2. This contrasts with the theories of Longuet-Higgins and Steedman (1971) and Steedman (1973, 1977) who propose that listeners listen hard at the beginning of a passage for cues that help them to induce a metrical structure and once they have induced a metre, it takes a lot of counterevidence to convince them to change to a different metrical structure.
3. Consider for example, this rhythmic sequence here [Play Lee46]. Now, I bet most of you heard it like this rather than like this. That is, you avoided hearing the second note as being a syncopation because you didn't have to. But this meant that you had to hear this later tone as being syncopated instead.
4. This demonstrates that the metrical structure you infer from a passage depends heavily on the fact that you build it up as the passage unfolds and not all in one go after you have heard the whole passage.

18. Lee's (1991) experiments

General conclusions:

1. Assumption that certain specific cues (e.g., dactyls) always trigger certain metrical structures is too simplistic.
2. The metrical structures that listeners infer are strongly dependent on the fact that music unfolds in time.

Specific conclusions:

1. Listeners assume downbeat falls on first note and second note is metrically strong.
2. Listeners revise this assumption if given counter-evidence before bar established as metrical unit.
3. Co-occurrence of major syncopation and weak long note is strong counter-evidence, long note on weak beat is weak counter-evidence.
4. Listeners prepared to revise metre to get a tactus between 300ms and 600ms.

18. Lee's (1991) experiments

1. Lee carried out three experiments designed to test the four basic assumptions that I've just been talking about, that is
    (a) that long notes are heard to begin on strong beats
    (b) that we like as many strong beats to occur at note-onsets as possible
    (c) that repetition is important in cueing metrical structure, and
    (d) that it is important to take into account the fact that listeners build up a perceived metrical structure as the music unfolds in time.
2. In all three experiments, experienced musicians were played various specially constructed rhythmic sequences and asked to notate the sequences using standard staff notation.
3. Lee drew the following general conclusions from the results of his experiments:
    (a) The hypothesis that certain cues (e.g., dactyl rhythms or long notes) always trigger certain metrical interpretations is too simplistic: as Lee puts it, "the more counter-evidence a particular interpretation contains, the more likely it is to be rejected in favour of an interpretation containing less counter-evidence" (Lee, 1991, p. 98).
    (b) the fact that listeners build up a perceived metrical structure as the music unfolds is very important in explaining why listeners infer the metrical structures that they do.
4. He also found, in particular, that:
    (a) listeners initially assume that the downbeat falls on the first note and that the second note is metrically strong;
    (b) listeners are prepared to revise this assumption if they find counter-evidence before the bar has been established as a metrical unit;
    (c) the co-occurrence of a major syncopation with a weak long note constitutes strong counter-evidence and the occurrence of a weak long note on its own weak counter-evidence; and

(d) listeners are prepared to revise their metrical hypothesis in the interest of obtaining a tactus within the range 300-600ms.

19. Lee's (1991) model

- describes new model that incorporates results of his experiments
- works in left-to-right fashion, comparing lengths of successive notes and considering their positions with respect to time of particular beats
- predicts time of first downbeat, length of bar and metre
- four ways in which Lee's (1991) model different from his earlier models and those of Longuet-Higgins and Steedman:
    1. new model capable of generating more than one metric structure for a given passage;
    2. assumes downbeat is on first note and only shifts it in light of counter-evidence;
    3. model can find metrical levels lower than first one it discovers;
    4. revises metrical structure to get tactus in range 300-600ms.

19. Lee's (1991) model

1. Lee describes a new model that is based on his earlier theories but takes into account the findings of his experiments.
2. Like the earlier models, it works in a left-to-right fashion, comparing the lengths of successive notes and considering their positions with respect to the times of particular beats.
3. The algorithm predicts the time of the first downbeat, the length of a bar and the metre.
4. Lee's (1991) model differed from his earlier models and those devised by Longuet-Higgins and Steedman in four important ways:
    (a) Unlike his earlier models and those of Longuet-Higgins and Steedman, Lee's 1991 model is capable of generating more than one possible metrical structure for a given musical passage. The model has a parameter which determines how 'tolerant' the model is to counter-evidence.
    (b) It assumes that the downbeat occurs on the first note of the sequence and only shifts it onto a later note in the face of counter-evidence against this assumption.
    (c) The model can find lower metrical levels than the first level that it discovers.
    (d) It can revise a metrical hypothesis in the interest of obtaining a tactus within the range 300-600ms.

20. Temperley's (2001) theory of metrical structure

- heavily influenced by GTTM;
- presents model of metre, phrasing, counterpoint, harmony, key and pitch-spelling;
- implemented as computer programs;
- optimisation problem solved using dynamic programming technique;
- evaluated models by comparing output with scores.

Temperley		GTTM
MWFR 1	=	MWFR 2
MWFR 2	=	MWFR 3
MPR 1	=	MPR 3
MPR 2	=	MPR 5a
MPR 3	≃	MWFR 4
MPR 4	≃	MPR 2
MPR 5	=	MPR 10
MPR 6	≃	MPR 5f
MPR 7	=	MPR 4
MPR 8		–
MPR 9	≃	MPR 1

20. Temperley's (2001) theory of metrical structure

1. I'm going to finish up by spending a few minutes discussing an important recent theory described by David Temperley in his book *The Cognition of Basic Musical Structures*, published in 2001.
2. In this book, Temperley presents a computational theory of music cognition that is deeply influenced by Lerdahl and Jackendoff's (1983). Like Lerdahl and Jackendoff, Temperley attempts to explain the cognition of common-practice music by means of a system that generates structural descriptions from musical 'surfaces'.
3. As in GTTM, the hypothesis underlying Temperley's theory is that the analysis it generates for a passage of music correctly describes certain aspects of how the passage is interpreted by listeners who are experienced in the idiom.
4. Like GTTM, Temperley's theory consists of a number of preference rule systems, each containing well-formedness rules that define a class of structural descriptions and preference rules that specify an optimal structural description for a given input. Temperley presents preference rule systems for six aspects of musical structure: metre, phrasing, counterpoint, harmony, key and pitch spelling.
5. In collaboration with Daniel Sleator, Temperley has implemented most of his theory as computer programs. Finding the best analysis satisfying a given set of preference rules is an optimisation problem that can be solved using a technique called dynamic programming (Bellman, 1957;Cormen *et al.*, 1990, Chapter 16). Each of Temperley's six models is implemented using the dynamic programming technique, the systems for harmony and pitch spelling being combined into a single program.
6. Temperley evaluated each of his six models using objective tests. For example, he tested his metre program on a corpus of 46 excerpts from a theory workbook by Kostka and Payne (1995), comparing the output of the program with the scores of the excerpts.
7. Most of the rules in Temperley's model of metre are borrowed from the metre component of GTTM. Table 1 shows how rules in Temperley's system correspond to rules in GTTM.
8. However, one important difference between Temperley's metre model and most previous approaches is that it can

Temperley		GTTM
MWFR 1	=	MWFR 2
MWFR 2	=	MWFR 3
MPR 1	=	MPR 3
MPR 2	=	MPR 5a
MPR 3	≃	MWFR 4
MPR 4	≃	MPR 2
MPR 5	=	MPR 10
MPR 6	≃	MPR 5f
MPR 7	=	MPR 4
MPR 8		–
MPR 9	≃	MPR 1

Table 1: Correspondence between rules in Temperley's theory of metrical structure and those in the metrical structure component of GTTM (see Lerdahl and Jackendoff, 1983, pp. 345–347 and Temperley, 2001, pp. 357–358). The symbol '=' indicates that the two rules are the same; the symbol '≃' indicates that Temperley's rule is a modification of the rule in GTTM; the symbol '–' indicates that there is no equivalent of Temperley's rule in GTTM.

process 'expressive' polyphonic input representations. All the theories I've been talking about so far only process melodies.

9. Lerdahl and Jackendoff's (1983, pp. 72, 347) MWFR 4 specifies that beats at the tactus level and above must be 'equally spaced'. Clearly, this only applies if the input is a 'metronomic' representation. Temperley enables his model to handle 'expressive' representations by re-expressing Lerdahl and Jackendoff's MWFR 4 as a preference rule (MPR 3), stating that beats at each level should be 'maximally evenly spaced' (Temperley, 2001, p. 35).
10. Also, Temperley includes a rule (MPR 8) that states that the metric structure should reflect the linguistic stress of any words that are being sung. There is no equivalent of this rule in LJ's theory.
11. In his computer implementation of this theory, he only incorporated the first 5 MPRs. When he ran his program on 46 excerpts from a harmony text book, he found that it correctly predicted the correct tactus in over 90% of cases when the input was quantized and around 85% of cases when the input was not quantized.

References

Bellman, R. (1957). *Dynamic Programming*. Princeton University Press.

Clarke, E. F. (1999). Rhythm and timing in music. In D. Deutsch, editor, *The Psychology of Music*, pages 473–500. Academic Press, San Diego.

Cormen, T. H., Leiserson, C. E., and Rivest, R. L. (1990). *Introduction to Algorithms*. M.I.T. Press, Cambridge, Mass.

Kostka, S. and Payne, D. (1995). *Workbook for Tonal Harmony*. McGraw-Hill, New York.

Lee, C. S. (1985). The rhythmic interpretation of simple musical sequences: Towards a perceptual model. In I. Cross, P. Howell, and R. West, editors, *Musical Structure and Cognition*. Academic Press, London.

Lee, C. S. (1991). The perception of metrical structure: Experimental evidence and a model. In P. Howell, R. West, and I. Cross, editors, *Representing Musical Structure*, pages 59–127. Academic Press, London.

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

Longuet-Higgins, H. C. and Lee, C. S. (1982). The perception of musical rhythms. *Perception*, **11**, 115–128.

Longuet-Higgins, H. C. and Lee, C. S. (1984). The rhythmic interpretation of monophonic music. *Music Perception*, **1**(4), 424–441. Republished as Longuet-Higgins and Lee (1987).

Longuet-Higgins, H. C. and Lee, C. S. (1987). The rhythmic interpretation of monophonic music. In H. C. Longuet-Higgins, editor, *Mental Processes: Studies in Cognitive Science*, pages 150–168. British Psychological Society/MIT Press, London, England and Cambridge, Mass. Published earlier as Longuet-Higgins and Lee (1984).

Longuet-Higgins, H. C. and Steedman, M. J. (1971). On interpreting bach. In B. Meltzer and D. Michie, editors, *Machine Intelligence*, volume 6, pages 221–241. University of Edinburgh Press. Republished as Longuet-Higgins and Steedman (1987).

Longuet-Higgins, H. C. and Steedman, M. J. (1987). On interpreting Bach. In H. C. Longuet-Higgins, editor, *Mental Processes: Studies in Cognitive Science*, pages 82–104. British Psychological Society/MIT Press, London, England and Cambridge, Mass. Published earlier as Longuet-Higgins and Steedman (1971).

Parncutt, R. (1994). A perceptual model of pulse salience and metrical accent in musical rhythms. *Music Perception*, **11**, 409–464.

Povel, D.-J. and Essens, P. (1985). Perception of temporal patterns. *Music Perception*, **2**, 411–440.

Steedman, M. J. (1973). *The Formal Description of Musical Perception*. Ph.D. thesis, University of Edinburgh.

Steedman, M. J. (1977). The perception of musical rhythm and metre. *Perception*, **6**, 555–569.

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