

Optimal Musical Grouping

Jonah Katz

I. Introduction

The Generative Theory of Tonal Music (Lerdahl & Jackendoff 1983) attempts to characterize human musical intuitions making use of *well-formedness rules (WFRs)* and *preference rules (PRs)*. Their formalism bears many similarities to constraint-based models of linguistic competence, such as Optimality Theory (Prince & Smolensky 1993). L&J's preference rules are essentially violable constraints; their well-formedness rules seem to be inviolable principles of structure-building.

In my research, I'm attempting to answer two questions:

1. Is there a way of implementing the GTTM preference rule system in an OT framework?
2. Does this reformulation have anything useful to offer the theory?

As a first pass, I've attempted to isolate *grouping structure*, the most basic of the four analytical components discussed in GTTM. For this component of the theory, I believe the answer to both of the above questions is "yes."

II. General Considerations

Optimality Theory is in some ways more explicit than the GTTM framework. In fact, the discussion of formalism in GTTM can be seen as defining a range of possible implementations for the preference rule system; an OT model, in turn, can be seen as one of those possible implementations. L&J purposely avoid giving a strict implementation of their preference system with weights and competition. This is a useful and understandable move, for it allows them to avoid getting bogged down in minute mathematical detail. It also allows for a number of ways of implementing the preference system. OT is one way.

In "classic" OT, constraints are ranked in a strict dominance hierarchy, and there is no possibility of additive constraint conjunction, or two low-ranked constraints "teaming up" to force violation of a high-ranked constraint. L&J don't give an explicit framework for evaluating the badness of one PR violation against another, but they clearly intend for additive effects to be possible.¹ Yet nowhere in the analyses they provide, at least for grouping structure, is such "teaming up" of constraints necessary. If we can do without teaming up, it will result in a more parsimonious theory of musical structure-building.

Nonetheless, some situation may pop up where gradient constraint weighting is necessary. At the very least, when discussing difficult judgements, it is sometimes the case that more than one analysis is plausible. It may be desirable in these cases to be able to evaluate well-formedness on a numerical scale. There are many versions of OT that relax the strict domination requirement. Gradient well-formedness or variation can be

¹ See pp. 46-48.

modeled using stochastic approaches to OT (Boersma 1998 et seq., Hayes 2000), while still ruling out “teaming up.” More complex teaming-up or trading-off relations can be modeled in fully additive frameworks such as Harmonic Grammar (Legendre et al. 1990) and various versions of additive OT (Flemming 2001, Pater 2006). In the discussion that follows, gradient well-formedness judgements and teaming-up effects will not be relevant. As such, I’ll model the grouping intuitions in “classical” strict-dominance OT. Any of these analyses are in principle translatable into a system with weighted constraints.

Importing the logic of the OT framework into GTTM allows a different kind of approach to basic structure-building. Grouping structure appears to take the form of a strictly-layered, exhaustively-partitioned hierarchy. In GTTM, this fact is the result of extensive “hard” or inviolable constraints on structure-building, known as the WFRs. In OT, the hierarchical nature of grouping structure falls out from the overall logic of the theory, with just a few foundational stipulations. This will be discussed later.

III. Grouping Structure

Grouping structure is the component of the analytical system that deals with how pitch events are parsed into a constituent structure. Listeners are able to extract various levels of constituency out of a musical surface, from a three- or four-note motif to a phrase or a major formal section. To take a simple example, consider the birthday song, shown in Figure 1.

Figure 1. The birthday song.



Anybody who knows this song knows certain facts about the constituent structure therein. For instance, the first six notes are in some sense parallel to the second six notes. They sound similar but not the same, and they each stand as a piece of music on their own. This is reinforced by the lyrics, where linguistic sentences correspond to this first level of grouping. Notes 2 through 4, for instance, don’t sound like a coherent piece when taken out of context, and the lyrics (“-py birthday”) do not form a linguistic constituent. The third group of seven notes starts out with the same motif as the first two groups (“happy” followed by an upwards leap to “birth-”) but ends differently, fragmenting or developing the original melodic idea. The fourth group is parallel to the first two, but ends in a cadence or structural ending. At a higher level, we naturally divide the piece into two parts of roughly equal duration. The break between these two parts is marked by a change in lyrics at the onset of the second half (“happy birthday to blah blah”). At the highest level, we know that this is one verse of the birthday song. If we went any further, we would be repeating this verse.

The goal of any theory of grouping structure is to develop a formal model of how these constituency intuitions may arise from a musical surface. In GTTM, the emphasis is on Gestalt principles from visual perception, such as proximity and change, and the ways in which they affect constituency. The idea is that grouping structure is fully determined by the relationship between pitch events in the musical surface; preference rules determine where grouping boundaries occur, and well-formedness rules determine how groups relate to one another structurally.

IV. Grouping is Not Meter

The existence of grouping as a separate component of musical cognition is not always obvious, especially to non-musicians. One might think that the impression of grouping is simply an emergent property of the fact that music has a rhythmic structure, and that notes occur with varying degrees of proximity in that rhythmic structure.

One argument that grouping exists is the fact that it can either be in agreement with the meter or can cut across it. One frequent kind of meter-grouping ‘clash’ is known as *hemiola*. Hemiola is when phenomenal accents are heard as being grouped in a meter that is different from the abstract metrical structure of the piece in question.²

Take as an example the Duke Ellington song “It Don’t Mean a Thing (if it Ain’t Got that Swing).” The second four bars of the form prominently feature hemiola and syncopation.

[Sound File]

There is a clearly a sense in this excerpt that musical events are cutting across the meter. This comes from a two note motif that repeats at intervals of one and a half beats. The meter, instead of one and a half beats, is in bars of four, which divide into two groups of two beats. The grouping of the pitch events, then, is repeating at a different rate than the strong beats of the meter. This is what gives us the feeling of “cutting across.”

One could imagine trying to reduce this phenomenon to syncopation, when phenomenal accents (pitch events) fail to coincide with rhythmic accents (strong beats). On this view, the mismatch between pitch events and the meter is a consequence of the fact that one out of every two notes in this excerpt occurs on an off-beat. But this type of an explanation can not capture this phenomenon. It predicts that the same cross-cutting intuition would still hold if the two-note motif were recurring at two-beat intervals, in phase with the rhythm:

[Sound file]

² Hemiola in its narrowest sense is limited to situations where pitch events are organized into three groups of two-beat duration when the underlying metrical structure is actually two measures of three beats each. I use the term in a looser sense common amongst musicians, to mean any situation where pitch events are grouped in a repeating rhythmic structure with a different number of beats than the meter of the piece.

The intuition here is clearly not one of cross-cutting, even though it is still the case that half of the notes are on off-beats. Grouping is the only way to explain this phenomenon. Metrical beats (rhythmic accents) are abstract; they occur whether or not they actually coincide with pitch events (phenomenal accents) in the auditory signal. The grouping of phenomenal accents can be either in-phase or out-of-phase with the grouping of rhythmic accents, and this is where the feeling of hemiola or cross-cutting comes from.

V. Optimality Theory

OT was originally developed as a logical framework in which linguists can construct models of human phonological competence. The basic idea is that the human language faculty provides some universal constraints on possible sound patterns in a language. For each specific language, there exists some ranking of the constraints that will generate that language. The input to this constraint system is the *underlying form*, some string of linguistic symbols at a certain level of abstraction. The output is a *surface form*, consisting of a more detailed phonetic representation of language-specific sound sequences. In order to avoid violating high-ranked constraints, we often produce output structures that violate lower-ranked constraints. The output form that incurs the fewest violations of higher-ranked constraints is the *optimal* candidate; it will be produced as output.

In adopting OT to grouping structure, we will take a musical surface as input, that is, a string of pitch events. Output candidates will consist of grouping structures (“analyses”) linked to that musical surface. The optimal analysis will violate the highest-ranked grouping constraints fewer times than all other possible analyses. Let us begin by looking at the ways in which preference rules about proximity and change can be restated as OT *surface disruption (SD) constraints*.

VI. Surface Disruptions

In GTTM, L&J make use of preference rules that insert group boundaries at moments of melodic, textural, dynamic, and temporal disruption in a musical surface. One rule, for instance, calls for a grouping boundary to be inserted at every break in phrasing in the musical surface. This calls for the grouping structure seen in Figure 2 (where the slur marks are taken to indicate phrasing).

Figure 2 (taken from GTTM). Break in phrasing.



These rules are strictly local; that is, they apply only to the first level of grouping structure. At higher levels of grouping, other rules take over. At an intermediate level, L&J introduce a rule that picks out “more pronounced” surface disruptions and calls for intermediate-level grouping boundaries. This means that there are at least three distinct classes of preference rule: one for the lowest level of grouping, one for intermediate

levels, and a third class of global rules that determine the highest levels of grouping structure (the most important one inserts a grouping boundary at important breaks in the time-span reduction; this handles, among other things, cadences).

In the OT framework, it is not necessary to distinguish between these three classes of rule. The SD constraints, for instance, will apply at all levels of grouping structure. At higher levels, however, other constraints (such as the cadence alignment constraint) will come into play, occasionally compelling violation of the lower-ranked SD constraints. The break in phrasing constraint, informally, will call for every break in phrasing in the musical surface to be aligned with the edge of some group *at every level of grouping structure*. More concretely, if any group at any level contains both of the pitch events flanking a break in phrasing, this constraint will be violated. Figure 3 shows two possible analyses of a musical surface, and compares their constraint violations in a *tableau*. The tableau is a table with output candidates (analyses) on the vertical axis and constraints on the horizontal axis. A violation is notated with an asterisk. As is clear, the analysis that groups across the break in phrasing (a) violates the constraint and the one that aligns a grouping break with the break in phrasing (b) does not. The exclamation point marks a *fatal violation*, one that eliminates an analysis from consideration. I'll call the break in phrasing constraint SD-Slur. The hand icon indicates the optimal candidate analysis.

Figure 3. An SD constraint at work.

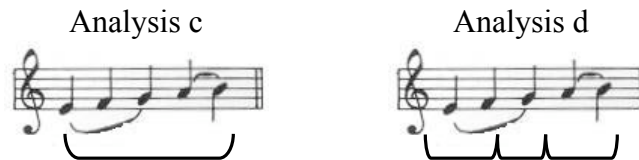


Figure 2	SD-SLUR
☞ Analysis a	
Analysis b	*!

The SD constraints ensure that grouping boundaries will occur at the lowest level of grouping structure. At the same time, they link those boundaries to disruptions in the musical surface; a break in grouping must have some reason for existing. To explain why we insert grouping boundaries *only* at points of disruption, and not arbitrarily at other points, we need a constraint *Structure. This constraint penalizes every single group present in an analysis. It will generally favor analyses that have fewer groups over those that have more. In GTTM, L&J posit a preference rule that favors larger groups over smaller ones; it does essentially the same work as this constraint.

*Struc is outranked by the SD constraints. This means that grouping boundaries can be inserted, but only to correspond to disruptions. In Figure 3.2 two more analyses (c and d) are added to the competition. Analysis c does not have any grouping breaks at all. Analysis d has too many. In this tableau, high-ranked constraints are on the left, low-ranked on the right. As can be seen, ranking *Struc below SD-Slur ensures that the optimal analysis A will be selected.

Figure 3.2. *Struc at work.



Figures 4, 10	SD-SLUR	*STRUC
☞ Analysis a		**
Analysis b	*!	**
Analysis c	*!	*
Analysis d		***!

VII. Higher-level Considerations

Because the SD constraints apply at all levels rather than just locally, many facts about intermediate-level grouping are already accounted for. In Figure 4 (taken from GTTM), three groups are present. Groups 1-2 are separated by registral leap (a third, as opposed to seconds within each group) and phrasing boundaries (slurs). Groups 2-3 are separated by a rest in addition to register and slur. In GTTM, we need an extra rule, and indeed an extra class of rules, to explain why a higher-level grouping boundary occurs in between 2-3. But in the OT framework, this fact falls out naturally from the SD constraints. Grouping across the 2-3 boundary at higher levels will violate three SD constraints (analysis b); grouping across 1-2 or 3-4 will violate only two (analysis a). Thus, the preferred binary parse comes out as optimal without any need to stipulate a division between local and global principles. In Figure 4, the piece level of grouping is included, but violations at this level are not included in the tableau; because the piece level is the same for both analyses, these violations can be said to cancel each other out. I've named the registral leap constraint SD-Register and the rest constraint SD-Pause. The dashed lines in the tableau indicate that no ranking is necessary between the SD constraints. No matter how we rank them, analysis a will always be more harmonic than analysis b. In OT, it is said that analysis a *harmonically bounds* analysis b; there is no reranking of these constraints that will allow analysis b to emerge as optimal.

Figure 4. Higher level grouping.

Analysis a



Analysis b

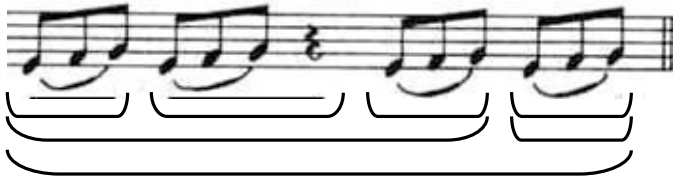


Figure 4	SD-SLUR	SD-REG	SD-PAUSE	*STRUC
☞ Analysis a	**	**	**	*****
Analysis b	**	**	***!	*****

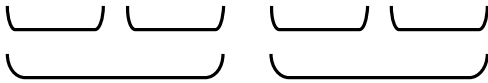
Certain other factors come into play in Figure 4. For one thing, it seems as if the optimal analysis would be no analysis at all. It wouldn't have any structure, so would not violate *Struc; it would not group across any surface disruptions, so would not violate SD constraints³. Clearly, this is an undesirable consequence of the theory in its current form.

VIII. Whence Structure?

The solution to the problem of why analyses have structure will require at least three components: we will need to explain why there is a piece level, why there is a lowest level, and why there are levels in between.

The solution to the piece level lies in one of L&J's well-formedness rules, alluded to above: every piece forms a group. This means that for any piece, there will be at least one level of grouping, that level containing the entire piece. This is a fundamental property of well-formed analyses; we can conceive of it as an undominated constraint. But perhaps a better way to think of it is as a logical necessity. One could imagine a grouping analysis that looks like this, with no piece level:

³ This holds if we assume that the SD constraints are formulated so as to penalize any group that dominates both of the pitch events flanking a disruption. It is conceivable that we could instead formulate the constraints to call for insertion of a grouping boundary at points of disruption. Depending on the precise formulation, this might result in violation when there is simply no analysis, as in the case we are considering here. But it would introduce a slew of other complications concerning the desirability of inserting a boundary at *some level* versus a boundary at *all levels*. I will not explore this alternative here.



But this is not an analysis of a piece of music; it is an analysis of two unrelated fragments of music. To say that a listener has formed a representation like this is a claim that the listener has not integrated the two halves of the excerpt into a coherent whole. To the extent that listeners hear an excerpt like Figure 4 as a piece of music, they have grouped it as a piece. Our theory doesn't have much to say about the case where they don't hear it as a piece of music.

The question of why there is a piece level, as noted above, is a sub-part of a larger question: why is there structure at all, as opposed to no structure? Any theory of grouping will have some type of economy constraint to ensure that an analysis doesn't contain any unnecessary structure. In the current approach, this is *Struc; in GTTM, it is the dispreference for small groups. In a derivational approach like GTTM, grouping structure is built from the bottom-up, and the preference rules call for the insertion of grouping boundaries at moments of disruption. To the extent that grouping differs from level to level, the preference rules must be different at higher levels. In a representational approach like OT, there is no such structure-building process. A musical surface stands in correspondence with some analyses, and the constraints call for certain properties to hold of that correspondence. The constraints, as formulated here, don't call for the insertion of anything; they simply state that breaks in the music are mirrored by breaks in the grouping. They are violable, and indeed all surface disruption constraints are violated at the piece level. Because they will all be violated at the piece level, it doesn't 'help' to obey them at lower levels; in fact, it makes the analysis worse, because it has more structure, hence more violations of *Struc. This is shown in Figure 4.2.

Figure 4.2. An undesired consequence.

Analysis c



Analysis d



Figure 4.2	SD-SLUR	SD-REG	SD-PAUSE	*STRUC
Analysis c	****	***	*	*
Analysis d	****	***	*	*****!

So why, in this theory, do we have constituents at all? The formal solution will need to stipulate this in Gen, the component of the system that generates analyses. It's simply ill-formed to generate candidates that group twelve notes together with no subdivisions. But this formal solution should ideally have an explanation; principles that are hard-wired into Gen should have some prior plausibility as universals of music cognition, or as domain-general cognitive principles.

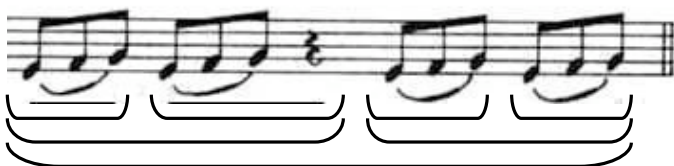
In this case, I suggest that the reason Gen subdivides pieces of music is *chunking* (Miller 1957). Human beings are simply incapable of forming mental representations of more than seven (plus or minus two) units without grouping them in some way. This constitutes a claim that analysis c above is not a possible mental representation, and as such is not a candidate in the current theory. Once we limit our attention to candidates that respect chunking, lower-level boundaries will become a logical necessity. The SD constraints will ensure that those boundaries fall in the right places; otherwise they would incur violations.

This is the second link in generating hierarchical structures ("unity of the piece" was the first link). Now we guarantee that every analysis has a highest level (the piece level). Any disruptions in the musical surface will induce a lower level or levels, because of chunking and the SD constraints. As discussed above, we also want intermediate levels, and the SD constraints can help with that. But it will also require a third link: the principle of binarity.

Without binarity, the optimal parse of figure 4 would be analysis c from Figures 4.2 and 4.3. This structure groups across each surface disruption once (at the piece level). It contains just five groups, one of which is the piece level. The optimal parse, as we saw above, should contain an intermediate level as well, as in analysis a. But this would group across two of the surface disruptions twice each, racking up more violations of SD constraints. And it would contain 7 groups, entailing more violations of *Struc. In fact, analysis c harmonically bounds analysis a with respect to our current constraint set. So how can our system generate such an analysis as optimal?

Figure 4.3. Another undesired consequence.

Analysis a



Analysis c



Figure 4.2	SD-SLUR	SD-REG	SD-PAUSE	*STRUC
Analysis a	*****!	*****	*	*****
Analysis c	***	***	*	*****

The constraint Binaricity says, “every group that immediately dominates another group must immediately dominate two groups.” Here *immediately dominates* means that group A dominates (>) group B, and there is no other group C such that A > C > B. Informally, this constraint says that every group contains two groups at the next level down. For now, the concept of a *level* of grouping doesn’t need to have any formal status, but it is a useful way to think about the hierarchy. This constraint exempts groups at the lowest level from the binarity requirement. The constraint only applies to groups that dominate other groups, and the lowest level of grouping dominates only notes (pitch events).

This is a very high-ranking constraint in Western tonal music. In order to favor optimal analysis B over the sub-optimal analysis A, Binaricity would need to be ranked above all of the constraints discussed so far. The competition would look like this:

Figure 4.2	BINARITY	SD-SLUR	SD-REG	SD-PAUSE	*STRUC
Analysis a		*****	*****	*	*****
Analysis c	*!	***	***	*	*****

Binaricity might also be considered as a component of Gen, an *a priori* principle of temporally-organized mental representations which is inviolable. The evidence for this formal move is not quite as strong as the evidence for chunking. Certainly, binarity seems to be a preferred feature of hierarchical mental representations. Strict binarity has been widely assumed for syntactic theory and elements of prosodic theory in linguistics. In music, the reductional components of GTTM are proposed with strict binary branchings. Meter, on the other hand, seems to have possibilities of being either double or triple. This is a counterexample to binarity as a putative universal. Some large-scale musical forms, such as sonata allegro, appear at first to be ternary. Sonata form is traditionally said to consist of exposition, development, and recapitulation sections. Yet even in this form, listeners often or always find some reason to group the development with either the exposition (in earlier versions of the form derived from rounded binary form) or the recapitulation (in later versions, especially during the Romantic period) as a subconstituent (Fred Lerdahl, pc). Thus, even this apparently ternary form features “hidden binarity.” It’s worth considering the possibility that perhaps all ternarity in language and music is instead recursive binarity. This is a very broad and general hypothesis, and I certainly can’t prove it here, so I’ll leave unresolved the question of

whether grouping binarity is a high-ranked violable constraint or an inviolable principle of cognitive structure-building.

IX. Interim summary

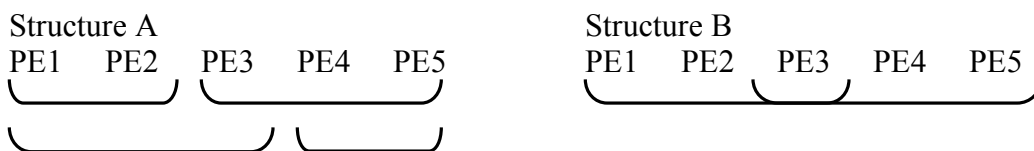
At this point, we've made a few rather major changes to the GTTM system. Rather than building up hierarchical grouping structure from two types of rule application, we've allowed the hierarchical structure to emerge from the interaction of three factors: disruptions in the musical surface for the lowest level of grouping (and to some extent intermediate ones), binarity for intermediate levels, and the "unity of the piece" for the highest level. With these assumptions, we generate most of the structures we'll need for grouping analysis.

Within the explicit OT ranking system, it is possible to examine in greater detail the ways in which constraints on analyses interact. In this system, the division into constraints on lower levels of grouping and constraints on higher levels of grouping is eliminated. Constraints are obeyed at global levels depending on how highly ranked they are. Constraints that seem to apply only locally are conceived of as low-ranked, but they can still exert an effect on higher levels of grouping. In figure 4, for instance, we saw how an SD constraint can exert a "tie-breaker" effect on higher levels of grouping. In the next section, we'll examine a few more issues in well-formedness.

X. Several Varieties of Partial Overlap

Many of the grouping WFRs deal with overlapping structures such as those in figure 5, where PE stands for pitch event.

Figure 5. Aberrant grouping structures.



In the current approach, the relative markedness of these structures falls out of the architecture of the theory, without the need for any additional well-formedness conditions on the structural relations that may hold between groups.

In structure A, a higher-level group contains part but not the entirety of a lower-level group, creating a mismatch between successive levels of grouping. In the current approach, such an analysis will never emerge as optimal. Because our SD constraints call for a grouping boundary at each level, all of the boundaries in structure A must be induced by the same surface disruptions. If we assume that disruptions occur in between events 2-3 and 3-4, there are two possibilities: either the disruption from 2-3 is more prominent (prominence being defined in terms of the SD constraint ranking), or the

disruption from 3-4 is more prominent. In either case, the optimal analysis will be one where the grouping breaks at all levels are aligned with the most prominent disruption. Any candidate aligning any break at any level with a less prominent disruption will be more marked than a candidate that respects the more prominent disruption. In the case where both disruptions are of equal prominence, our theory correctly predicts that no candidate will emerge as optimal; more information is needed in this case. L&J discuss this in GTTM (ch. 3), and conclude that the rules should not apply in this ambiguous case. They deal with it by formulating the proximity rules relative to a four-note context, a complication which is not necessary in the OT approach.⁴

Structure B, on the other hand, can occur under special circumstances in the current approach. Under normal circumstances, it will not emerge as optimal. Assuming again prominent disruptions at 2-3 and 3-4, note that each group crosses one disruption once. An exhaustive, non-overlapping grouping layer would only cross one of the two disruptions, so it harmonically bounds this candidate analysis. Note, however, that structure B is optimal in one class of situations, referred to as grouping *overlap* and *elision*. In overlap, a single chord (usually a tonic) serves as both a structural ending to a preceding group (a cadence, essentially) and a structural beginning to a succeeding group. In this case, the doubly-functioning pitch event will be contained in both groups, as in structure B. Here is an example from the children’s song *Miss Lucy*:

Figure 6. Grouping overlap in *Miss Lucy*.



The note labeled “Ov.” here serves as the structural ending of one verse and the structural beginning of the next. The lyrics are composed to exploit this overlap in function. Linguistic overlap co-occurs with musical overlap, disguising naughty words:

*Miss Lucy had a steamboat, the steamboat had a bell
Miss Lucy went to heaven, the steamboat went to HELL
(HELL)-o operator, please give me number nine
And if you disconnect me, I’ll chop off your BEHIND
(BEHIND) the ‘frigerator, there was...*

⁴ Except in the case of “registral leap,” a concept that must be defined as a distance relation greater than all local distance relations.

We will need two constraints, one anchoring every structural ending to the right edge of some group at every level of structure, and one anchoring every structural beginning to the left edge.⁵ These constraints will force the relatively marked structure referred to as structure B above. Structure B will never occur in the absence of some event that needs to be anchored to the left edge of some group and the right edge of some other group.

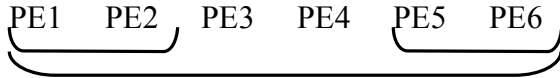
In GTTM, grouping overlap is dealt with by splitting grouping structure into two components: an underlying grouping structure, in which the ambiguous pitch event is actually two events, and a surface structure, in which they fuse. This necessitates an entirely new level of underlying representation, a well-formedness condition prohibiting structure B, and a new class of transformational rules to proceed from underlying to surface representations. Most of this is unnecessary in the current theory. Because structure B is independently more marked than a comparable non-overlapping structure, it will never occur without a compelling reason. That compelling reason is given by situations where one event simultaneously serves two functions, each associated with a particular position in grouping structure. I take from GTTM the crucial insight that one pitch event is occupying two positions in grouping structure. Within the current theory, there is a natural way to derive this, because there is no well-formedness condition militating against such a structure. This is yet another advantage of a representational system over a derivational one.

One other ill-formed structure is not yet excluded by the OT model. It is a non-exhaustive parse of the type seen in figure 6:

Figure 6. Another ill-formed structure.

⁵ It is not yet entirely clear how these structural accents will be defined. In GTTM, they are basically primitives in time-span reduction, and most of the “work” of these constraints is accomplished by interface conditions between grouping and time-span reduction. For now, I’ll also take structural accents to be primitives in the theory. In the future, it will be desirable to see if we can define them by reference to one of the reduction components. Indeed, Fred Lerdahl (pc) suggests that a cadence is a learned lexical item, essentially a primitive, but that a structural beginning can be defined as the most stable event in a cadenced group which precedes the cadence. Stability itself is discussed brilliantly and at length in Lerdahl 2001.

An entirely different kind of approach would seek to derive the overlap analysis here from considerations of parallelism. The end of the verse, with a cadence, is clearly parallel to the end of the first four-measure phrase, which contains a half-cadence. At the same time, the overlap note is also clearly parallel to the beginning of the piece; in fact, it begins a group which is identical to the beginning of the piece. It would be possible to derive this analysis, then, with a parallelism constraint that calls for parallel parts of the musical surface to be anchored to parallel positions in grouping structure. This approach would work for the current analysis, and would eliminate the need to refer to structural accents as primitives. But the issue of how to define parallelism is in itself quite difficult. For the current paper, I’ll simply note this possibility and move on.



Here, the largest group immediately dominates two smaller groups and a sequence of pitch events occurring in between the two groups. Note that this structure satisfies binarity. It also satisfies the SD constraints. Assuming disruptions in between PE2-3 and PE4-5, this parse does not group across any disruptions at the lowest level. It also has fewer total groups than the desired exhaustive parse, meaning fewer violations of *Struc. Note that we could change the binarity constraint to rule out this structure, by saying that a group should immediately dominate two groups *and nothing else*. But this seems to me to conflate two phenomena: binarity is a consideration that may in principle be violated, but the type of analysis in figure 6 seems incoherent, and the theory should never allow it to occur.

Intuitively, this structure is ill-formed because it mixes and matches groups and ungrouped pitch events. It's hard to know how we would even interpret this structure; I suggest that it should be banned from the theory outright, as a condition on the Gen component that generates candidate analyses. The condition can be stated as "a group that immediately dominates a group immediately dominates only groups." We'll "hard-wire" this into the model, so that this class of analyses is never entertained. At the same time, hard-wiring it is an admission that we can't explain it. Again, I will assume that this is a universal feature of temporally-organized cognitive representations; perhaps groups and pitch events are just too different to bear any grouping relationship to one another except for containment.

XI. Summary

I've shown how adapting the Optimality Theory framework allows us to preserve the many insightful generalizations of GTTM in an explicit and parsimonious form. Several of the well-formedness principles that needed to be specified as inviolable rules in GTTM fall naturally out of the architecture of OT. The nested hierarchical structure of grouping comes from an interaction between constraints calling for a piece to be a coherent whole, calling for grouping structure to reflect disruption in the musical surface, and calling for all grouping dominance relations to be binary. In this system, a nested hierarchical structure is simply the best way to satisfy all of these constraints at once. Several other ill-formed structures are ruled out because in the current model, they are competing with well-formed structures that are always more harmonic. This allows us to significantly reduce the number of well-formedness conditions in the theory.

The distinction between local details and higher-level principles implicit in GTTM has been eliminated in the OT model. The fact that higher-level considerations such as binarity often seem to trump local detail rules is a reflection of the constraint ranking. All constraints apply at all levels. The representational approach means that principles of grouping are the same at all levels, differing only in the size of the groups under question. In a derivational approach, by contrast, one might imagine the groups at one level being constructed according to entirely different principles than the groups at another level. To

the extent that there is no evidence for such disparities, they should be ruled out of the theory.

Finally, the need for underlying and surface structures is eliminated in an OT model. In this model, overlapping of groups is induced only by compelling evidence from structural accents. The fact that overlap does not occur in other positions is a reflection of its marked status relative to non-overlapping grouping structure. A structurally ambiguous accent is essentially the only situation where grouping overlap is a viable analysis.

The success of the OT approach in characterizing grouping structure suggests that it will be desirable to approach other components of the theory in the same framework. The reductional components, in particular, are likely to interact with grouping structure in interesting ways. It remains to be seen if that interaction will prove as amenable to an OT model as the grouping component seems to be.

References

- Boersma, P. 1998. Functional phonology: Formalizing the interactions between articulatory and perceptual drives. Ph.D. dissertation, University of Amsterdam.
- Flemming, E. 2001. Scalar and categorical phenomena in a unified model of phonetics and phonology. *Phonology* 18: 7-44.
- Hayes, B. 2000. Gradient well-formedness in Optimality Theory. In J. Dekkers, F. van der Leeuw, & J. van de Weijer (eds.), *Optimality Theory: Phonology, Syntax, and Acquisition*. Oxford, UK: Oxford University Press. 88-120.
- Legendre, G., Y. Miyata, & P. Smolensky. 1990. Harmonic grammar: A formal multi-level connectionist theory of linguistic well-formedness: Theoretical foundations. Technical Report 90-5, Institute of Cognitive Science, Univ. of Colorado.
- Lerdahl, F. & Jackendoff, R. 1983. *A Generative Theory of Tonal Music*. Cambridge, Mass.: MIT Press
- Lerdahl, F. (2001). *Tonal Pitch Space*. New York: Oxford University Press.
- McCarthy, J. 2002. *A Thematic Guide to Optimality Theory*. Cambridge, UK: Cambridge University Press.
- Pater, J. 2006. Additive optimization and phonological typology. Handout from the UMass-MIT Meeting in Phonology. University of Massachusetts Amherst, February 2006.
- Prince, A. & P. Smolensky. 1993. *Optimality Theory: Constraint Interaction in Generative Grammar*. Report no. RuCCS-TR-2. New Brunswick, NJ: Rutgers University Center for Cognitive Science.

