An Overview of Hierarchical Structure in Music

Author(s): Fred Lerdahl and Ray Jackendoff

Source: *Music Perception: An Interdisciplinary Journal*, Winter, 1983/1984, Vol. 1, No. 2, Hierarchical Structure in Music (Winter, 1983/1984), pp. 229-252

Published by: University of California Press

Stable URL: https://www.jstor.org/stable/40285257

Music Perception
Winter 1983–84, Vol. 1, No. 2, 229–252

An Overview of Hierarchical Structure in Music

FRED LERDAHL
Columbia University
RAY JACKENDOFF
Brandeis University

After reviewing their theoretical approach, the authors present four kinds of hierarchical structure in music: grouping structure, metrical structure, time-span reduction, and prolongational reduction. A musical analysis is given to illustrate these structures. In the ensuing discussion, comparison is made with Schenker's theory, the nonoverlapping condition on hierarchies is considered, and relationships among the hierarchical, associational, and implicative dimensions of musical structure are mentioned.

In our book, *A generative theory of tonal music* (henceforth *GTTM*), we propose a detailed theory of musical hierarchies.[1] In this paper we (1) sketch our theoretical approach, (2) present some essentials of the theory through an analysis of a Bach chorale, and (3) discuss some general questions about musical hierarchies arising from the analysis.

Theoretical Perspective

GTTM develops a grammar of tonal music based in part on the goals, though not the content, of generative linguistics. The grammar is intended to model musical intuition. It takes the form of explicit rules that assign, or "generate," heard structures from musical surfaces. By "musical surface" we mean, broadly, the physical signal of a piece when it is played. By "heard structure" we mean all the structure a listener unconsciously infers when he

Fred Lerdahl is a composer and associate professor of music at Columbia University. Ray Jackendoff is professor of linguistics and chair of the Program in Linguistics and Cognitive Science at Brandeis University. In addition to his work with Fred Lerdahl on music theory, he is the author of *Semantic interpretation in generative grammar*, *X-Bar syntax*, and *Semantics and cognition*.
Requests for reprints may be sent to Fred Lerdahl, Department of Music, Columbia University, New York, New York 10027, or Ray Jackendoff, Department of English, Brandeis University, Waltham, Massachusetts 02154.

listens to and understands a piece, above and beyond the data of the physical signal.

There are two related ways in which the theory differs from previous music theories. First, it is psychological, in that it attempts to explicate a cognitive capacity. Listeners hear certain structures rather than others. How can these structures be characterized, and by what principles does the listener arrive at them? One would ultimately hope to specify those cognitive principles, or "universals," that underlie all musical listening, regardless of musical style or acculturation. Second, the theory attempts to produce formal descriptions in a scientific sense. That is, the goal is not just the description of formal relations, as happens in mathematics and in certain recent varieties of music theory. Rather, the descriptions pertain to something in the "real" world—even though, in this case, the reality is mental.

Thus the theory is predictive. In addition to criteria of internal coherence and parsimony, its principles can be verified or falsified by comparing the analyses it generates with one's intuitions about particular pieces of music. In addition, many of its principles can be investigated through laboratory experiment.

Implicit in this program of research are two simplifying idealizations familiar to cognitive science. First, we assume an "experienced listener." Obviously, no two listeners are exactly alike, nor are any two listenings by the same listener. But once a listener is familiar with a musical idiom, he is highly constrained in the ways he hears a piece in the idiom. A theory of musical understanding needs to characterize these common constraints as a foundation for the study of individual differences in hearing. Second, the theory provides structural descriptions only for the final state of a listener's understanding of a piece. In our judgment, a substantive theory of real-time listening processes cannot be constructed without first considering what information these processes must deliver.[2]

The theory is at present restricted in scope in two important ways. First, it focuses on classical tonal music. One cannot hope to address in any deep way the question of musical universals without first developing a precise

1. Lerdahl and Jackendoff (1983). Also see Lerdahl and Jackendoff (1977, 1981), Jackendoff and Lerdahl (1981, 1982).

2. This idealization corresponds to one sense of Chomsky's (1965) notion of "competence." It is also parallel to Marr's (1981) pursuit of a "computational" level of description in the theory of vision. Marr shows how nearly every previous theory of visual processing has failed because of inadequate attention to this aspect of the problem. Along with Chomsky and Marr, we do not disparage theories of real-time processing; they are an essential part of a complete psychological theory. But, methodologically, it appears crucial to characterize mental structures before asking how they are computed over time (see Jackendoff, in press, for more extended remarks).

theory of at least one complex musical idiom. (However, evidence toward the analysis of other tonal idioms appears at a number of points in *GTTM*.) Second, the present form of the grammar deals explicitly with only those aspects of heard structure that are hierarchical.

By *hierarchy* we mean an organization composed of discrete elements (or regions) related in such a way that one element may subsume or contain other elements. The elements cannot overlap; at any given hierarchical level the elements must be adjacent; and the relation of subsuming or containing can continue recursively from level to level.

Our theory proposes that four types of hierarchical structure are associated with a tonal piece. *Grouping structure* describes the listener's segmentation of the music into units of various sizes. *Metrical structure* describes the hierarchy of beats that he attributes to the music. *Time-span reduction* establishes the relative structural importance of pitch-events within the heard rhythmic units of a piece. *Prolongational reduction* develops a hierarchy of pitch stability in terms of perceived patterns of tension and relaxation. (This last component is the closest equivalent in our theory to Schenkerian reduction.)

Each of these structures is described formally by a separate component of the musical grammar, and within each component there are three rule types. *Well-formedness rules* (WFRs) provide the conditions for hierarchical structure for each component. *Transformational rules* (TRs) permit a constrained class of modifications on musical surfaces so that certain apparently ill-formed phenomena (such as grouping overlap and elision) can be treated as well-formed. *Preference rules* (PRs) establish which formally possible structures correspond to the listener's actual hearing of a given piece. Thus WFRs and TRs describe formal conditions, and PRs relate formal conditions to particular musical surfaces.

It is impossible within the space of this paper to explain the actual rules. But a few words should be included here about the nature of the PRs, since they are an innovation as a rule-type within generative grammars. These rules, which do the major work of analysis within the theory, pick out features in musical passages that influence the listener's intuitions. In the grouping component, for example, one grouping PR marks a potential grouping boundary at a pause in the music; another detects thematic parallelism between two groups which are potentially far apart at the musical surface; a third encodes the effect of large-scale pitch structure on grouping decisions. As these instances suggest, some PRs are local in application, others are global, and some relate effects across the four components. Out of this process emerges the most "preferred," or most coherent, analysis or analyses for the piece in question. Generally, a musical passage in which the various rules are mutually reinforcing strikes the listener as clear or stereotypical. Where the rules conflict, on the other hand, the musical structure seems vague or ambiguous, and more than one overall structural description may be assigned.

Thus the grammar marks not the categorical correctness or incorrectness of an analysis but rather its relative viability. Although this characteristic may seem unusual by comparison with standard linguistic grammars, it is quite normal in theories of vision (see Koffka, 1935 and Marr, 1981, for example) and appears to be ubiquitous in cognitive systems (see *GTTM*, and Jackendoff, in press).

An Analysis

We turn now to a rule-generated analysis of the Bach chorale, "Ich bin's, ich sollte büssen" (from the *St. Matthew Passion*). We have chosen this piece because it is short yet musically rich, and is easy to play at the piano; moreover, some readers may wish to make a comparison with Schenker's (1932) analysis. Our explanations will be brief; the reader is urged to follow the intuitive sense of the examples. For those interested in pursuing more closely the workings of the grammar, we will notate in parentheses the rules of *GTTM* that apply most critically.

Grouping Structure

When hearing a musical surface, listeners chunk it into motives, phrases, and sections. We represent groups by slurs placed beneath the music in an embedded fashion. Figure 1 gives the music of "Ich bin's . . ." together with its grouping analysis.

As with most Bach chorales, the grouping of "Ich bin's . . ." is simple and unambiguous. There are no salient groupings beneath the phrase level. The phrase endings are marked by fermatas (GPR 2a and perhaps GPR 2b, depending on whether the fermatas are interpreted just as indicators for breathing or also as actual "holds") and by cadences (GPR 7 in conjunction with TSRPR 7). At larger levels the phrases are organized by a high degree of symmetry and parallelism (GPRs 5 and 6). In particular, the second half of the piece is melodically almost identical to the first half: phrases *a* and *b* are repeated in phrases *d* and *e*, and phrase *f* completes phrase *c*. These considerations group together *a* and *b* into *g*, and *d* and *e* into *h*; similarly, *g* and *c* group into *i*, and *h* and *f* into *j*. Finally, the whole piece is perceived as a group (group *k*) (GWFR 2).

Metrical Structure

Insofar as the signal permits, the listener infers from a musical surface a hierarchy of strong and weak beats. In the classical tonal idiom, beats are

Figure 1.

Beat:	6	1	2	3	4	5	6	1	2	3	4	5	
𝅘𝅥𝅮 level	.	.	.	.	.	.	.	.	.	.	.	.	(etc.)
𝅘𝅥. level		.			.			.			.		
𝅗𝅥. level		.						.					

Figure 2.

equidistant, and strong beats occur every two or three beats apart at any given metrical level. If a beat at a particular level is felt to be strong, it is also a beat at the next larger level; this is how metrical structure is hierarchical. The notated meter, which of course is not heard as such but is a visual cue for the performer, usually indicates an intermediate metrical level.

Because beats do not have duration but are points in time, we notate metrical structure by rows of dots, as in Figure 2. Here beats 2, 3, 5, and 6 are weak, and are beats only at the eighth-note level; beat 4 is stronger, and, along with beat 1, is also a beat at the dotted quarter-note level; beat 1 is the strongest, and is also a beat at the dotted half-note level.[3]

The metrical structure of "Ich bin's . . .", shown in Figure 3 for the first phrase only, derives in a straightforward fashion. The phrases are of equal length and exhibit considerable parallelism, so the derivation for the first phrase suffices for the whole piece (MPR 1). Here the attack patterns and the regular harmonic rhythm clearly establish the quarter-note level (MPRs 3 and 5f). At the half-note level there is a momentary ambiguity due to the occurrence of the identical tonic sonority three and then four beats apart (MPR 9): it is initially unclear that the opening tonic is to be heard as an upbeat. But the prolongation of both the bass and the harmony in the second half of measure 1 (MPRs 5e and 6, MPR 5f), plus the long note in the melody and the suspension at the beginning of measure 2 (MPRs 5a and 8), all conspire to establish the half-note level as notated. The long melodic note and the suspension also support the given metrical structure at the whole-note level. This choice is reinforced by the general preference for strong beats relatively early in a phrase (MPR 2); the other possible choice for the whole-note level would lead to an anacrusis three quarters long, a less stable situation.

3. This notation is preferable to the traditional prosodic notation (– ◡), which obscures the relationship between strength of beat and hierarchical level. In Cooper and Meyer (1960), matters are further complicated by the intermingling of prosodic notation with the properties of grouping structure (see *GTTM* for further discussion).

Figure 3.

To the analysis in Figure 3, Figure 4 adds the *time-span segmentation*, which describes the apprehended rhythmic units produced by the interaction of grouping and meter. Weak beats are bracketed with the previous strong beats as *afterbeats*, unless grouping boundaries intervene, in which case they are bracketed as *upbeats* to the following strong beats. Besides capturing the distinction between upbeats and afterbeats, the time-span segmentation serves as input to the time-span reduction, as will be seen shortly.

Reductions

A reduction in music theory is a way to represent hierarchical relationships among pitches in a piece. Pitches perceived as relatively embellishing can be "reduced out" recursively, leaving at each stage a simplified residue of structurally more important material. At the end of this process only one event remains—the most stable structure, or *tonic*. The term "tonal" can be broadly defined as referring to music that is heard in such a hierarchical fashion.

In our theory we have tightened the notion of reduction by adding the following conditions: (a) pitch-events are heard in a strict (nonoverlapping) hierarchy; (b) structurally less important events are heard as elaborations of specific more important events, rather than simply as insertions or interpolations between more important events. Time-span reduction and prolongational reduction, while expressing different musical intuitions, both fulfill these conditions.

We notate reductions by means of trees. A right branch such as *e2* in Figure 5 signifies an event that is subordinate to a preceding event; a left

Figure 4.

branch such as *e3* signifies an event that is subordinate to a succeeding event. More specifically, *e2* in Figure 5 is subordinate to (or embellishes) *e1*, *e3* is subordinate to *e4*, and *e4* is subordinate to *e1*. Thus there is an underlying level in which *e2* and *e3* have been reduced out and in which *e4* is adjacent to *e1*.

Unlike syntactic trees in linguistic theory, reductional trees do not express grammatical categories. Linguistic trees represent *is–a* relations: an adjective plus a noun is a noun phrase, a verb plus a noun phrase is a verb phrase, and so forth. A grammatical category combines with another to form a third grammatical category. Musical trees, by contrast, represent *elaborational* relations: *e3* (Figure 5) elaborates *e4*, *e4* elaborates *e1*, and so forth. An elaborated event does not disappear or change from level to level, but proceeds intact up the tree until it too is an elaboration. The absence of grammatical categories in music marks a profound difference between music and language.[4]

4. The elaborational nature of pitch hierarchies and the absence of grammatical categories in music are confirmed by more than 400 years of music theory. For support of this view from a psychological standpoint, see Deutsch and Feroe (1981). On the other hand, see Keiler (1977) for a music-theoretic approach based on quasi-linguistic trees.

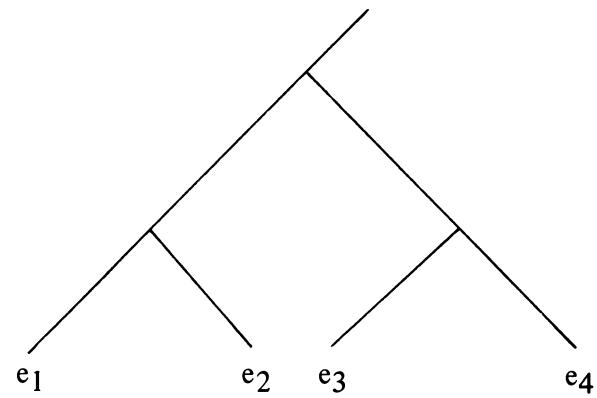


Figure 5.

Time-span Reduction

Time-span reduction expresses the way in which pitch-events are heard in the context of hierarchically organized rhythmic units. As suggested by Figure 4, these units, or *time-spans*, are determined by metrical structure at the most local levels, by a combination of meter and grouping at intermediate levels, and by grouping structure at large levels. Within each unit, events are compared in terms of mutual "stability," depending on such factors as the metrical position of the events, their intervallic configurations, voice-leading properties, and so forth. Starting at the most local levels, the most stable event in a unit is selected as *head*; other events in the unit are elaborations of the head, and are reduced out. Each head then goes on to the next larger time-span for comparison against another head from another time-span; and so on, until a single event is head for the whole piece.

Figure 6 illustrates for the first phrase of the chorale. Since the tree notation is unfamiliar, we include in the example a more traditional reductional notation beneath the music. The correspondence between the two notations is indicated by the labeling of the time-spans by level, of the nodes in the tree, and of the levels in the musical notation. For instance, level *b* in the musical notation consists of the heads of the time-spans labeled *b*. In the tree, this level consists of the branches that pass through or terminate at the node labeled *b*, leaving out all the branches that terminate lower. The choices for head in this passage are determined chiefly by relative local consonance (TSRPR 2), supplemented by metrical position (TSRPR 1).

To proceed further with the reduction, we must introduce the notions of *structural beginning* and *structural ending* (or *cadence*), which articulate grouping structure at the phrase level and all larger levels. The essential motion of a phrase takes place between these two structural points. A

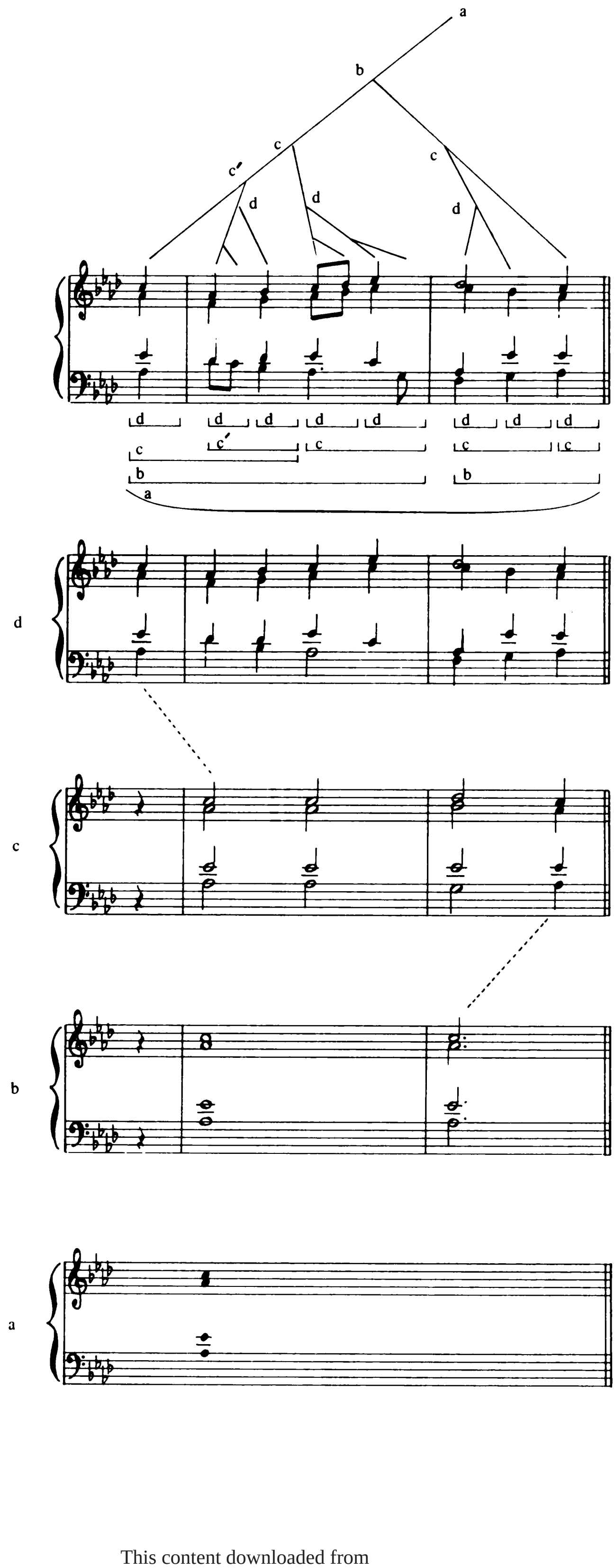


Figure 6.

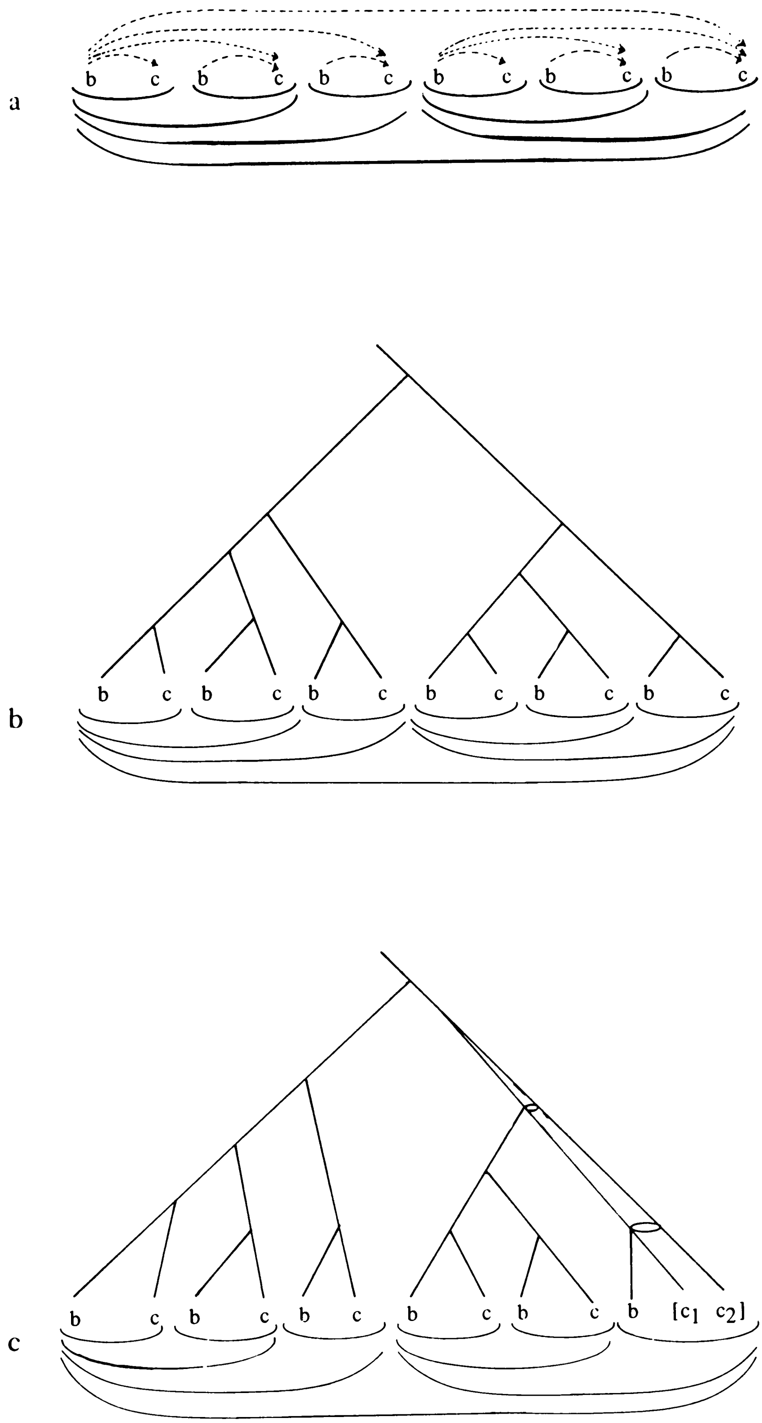


Figure 7.

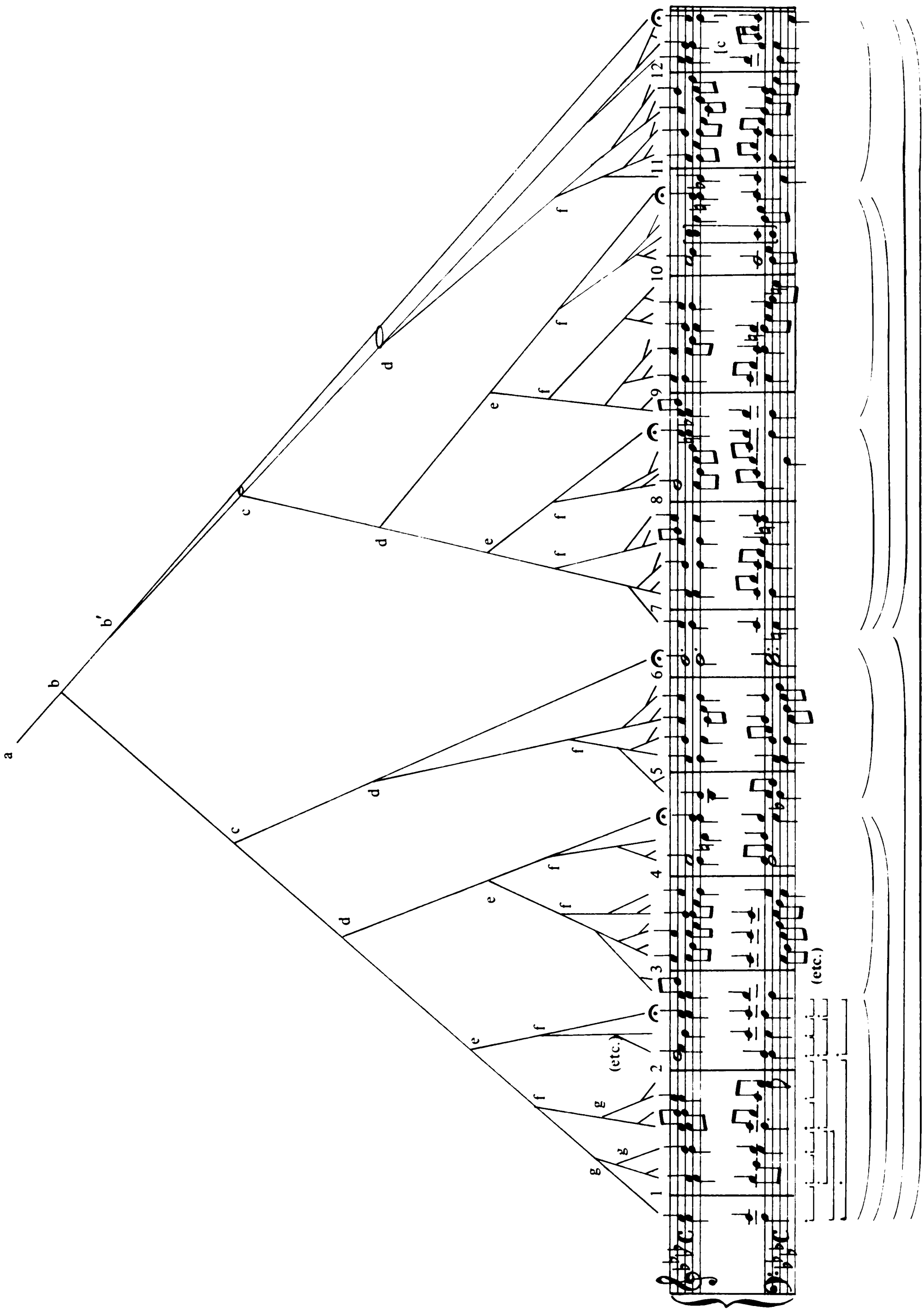


Figure 8.

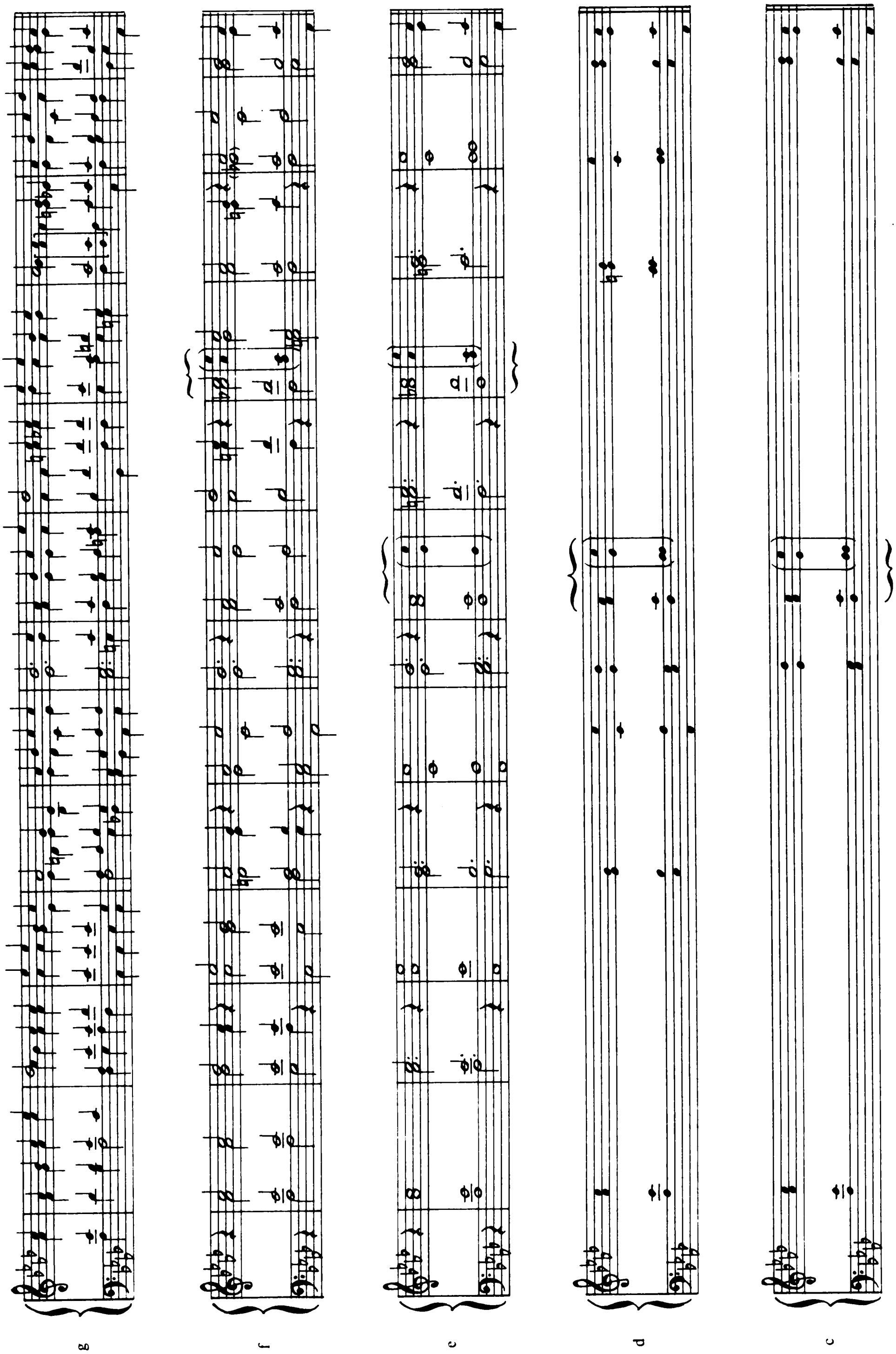


Figure 8. (continued)

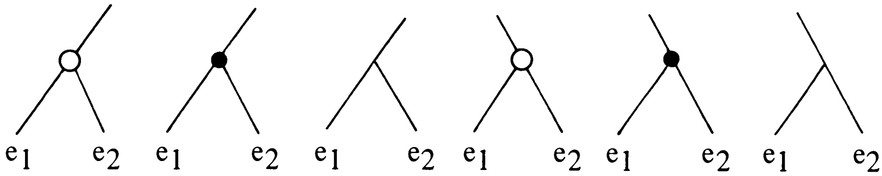
structural beginning is a relatively stable event appearing early in the phrase. A cadence is a harmonic/melodic formula of one or two members (in tonal music, most often $\overset{\hat{2}}{\mathrm{V}}$ or $\overset{\hat{2}}{\mathrm{V}}$–$\overset{\hat{1}}{\mathrm{I}}$) at the end of the phrase. The structural importance of a particular structural beginning or cadence to the piece as a whole depends on its position in the larger grouping structure: if it abuts a boundary of a large-scale group, it serves as a structural beginning or cadence for that group (TSRPRs 7 and 8). Figure 7a shows schematically the hierarchical function of each structural beginning (marked *b* in the diagram) and cadence (marked *c* in the diagram) in the grouping structure of "Ich bin's . . . ." Figure 7b translates Figure 7a into the tree notation. Figure 7c adds the reduction of the two-membered final cadence (the egg-like shapes in the tree signify subordination to the cadence as a whole).

Figure 8 completes Figure 6 and Figure 7 by giving the time-span reduction for the entire chorale. Levels *a–e* in Figure 8 correspond to Figure 7c; that is, these levels show the global hierarchical relationships among structural beginnings and cadences. Starting with level *f*, the branches represent local elaborations within the individual phrases. A good way to penetrate the analysis is to play at the piano the actual piece and then its successive reductional levels as represented in the musical notation. Each level should sound like a natural simplification of the previous level.

The naturalness of the reduction in Figure 8 is strained in the fourth and fifth phrases (measures 7–10), due to a structural ambiguity that requires comment. The melody in these measures repeats that of the first two phrases, and implies the tonic region. But the harmonization establishes the relative minor region, while at the same time ingeniously passing through the tonic in the course of each phrase (measure 7 and measure 9). Are these phrases, then, in the relative minor, or are they in the tonic with deviant cadences? As with Wittgenstein's (1953) rabbit–duck figure, a single response is insufficient. The musical grammar mirrors this situation through conflicting rule applications: it is a question of local harmonic stability (TSRPR 2b) versus a larger parallelism (TSRPR 4). In the tree we have opted for the relative-minor interpretation, but have indicated the alternative in brackets in the relevant levels of the musical notation.

Prolongational Reduction

One of the most important kinds of intuition a listener has is how the motion among the pitch-events of a piece tenses or relaxes. The prolongational component is intended to express such intuition explicitly and in detail. Briefly, if two events are heard as connecting prolongationally (at the musical surface or at an underlying level), and if the second event is heard as less stable than the first, the overall progression is felt as tensing. If the



a b c d e f

Figure 9.

second event is more stable than the first, the progression is felt as relaxing. Further, the degree of tension or relaxation between two events is determined by the degree of continuity between them. In a succession of two identical events, there is little sense of tension or relaxation. If an event is followed by itself in another form (say, by a chordal inversion), there is some sense of tension or relaxation (one or the other, depending on the relative stability of the two events). If an event is followed by a completely different event, there is a stronger sense of tension or relaxation.

In the prolongational tree, we represent a tensing pattern by right branching and a relaxing pattern by left branching. Figure 9a–c gives right-branching patterns, 9d–f left-branching patterns. The various nodes in the trees indicate contrasting degrees of continuity, hence of tension and relaxation. An open circle, as Figure 9a and 9d, describes an attachment where an event is replicated (say, $\overset{\hat{3}}{\mathrm{I}}$–$\overset{\hat{3}}{\mathrm{I}}$); this is called *strong prolongation.* A filled-in circle, as in Figure 9b and 9e, describes an attachment where an event progresses to another form of itself (say, I–I^6 or $\overset{\hat{3}}{\mathrm{I}}$–$\overset{\hat{1}}{\mathrm{I}}$); this is called *weak prolongation.* A node without any circle, as in Figure 9c and 9f, describes an attachment where an event progresses to a completely different event (say, I–V or ii^6–V); this is called *progression.*

Figure 10 illustrates these branchings through a prolongational analysis of the first phrase of "Ich bin's . . . ." As an aid, we again supply a secondary notation beneath the music; the slurs stand for branching connections, with the dashed slurs reserved for strong prolongations. The tree twice shows a tensing–relaxing pattern—from the opening sonority to its prolongation on the third beat of the first measure, and similarly from there to the end of the phrase.

We have found that a particular pattern of tensing and relaxing is needed if the listener is to experience *closure*: there must be at least one right-branching progression (9c) in the course of a group, followed by two left-branching progressions (9f) at the end of the group (two because the first left branch must be "prepared"). We call such a pattern *normative prolon-*

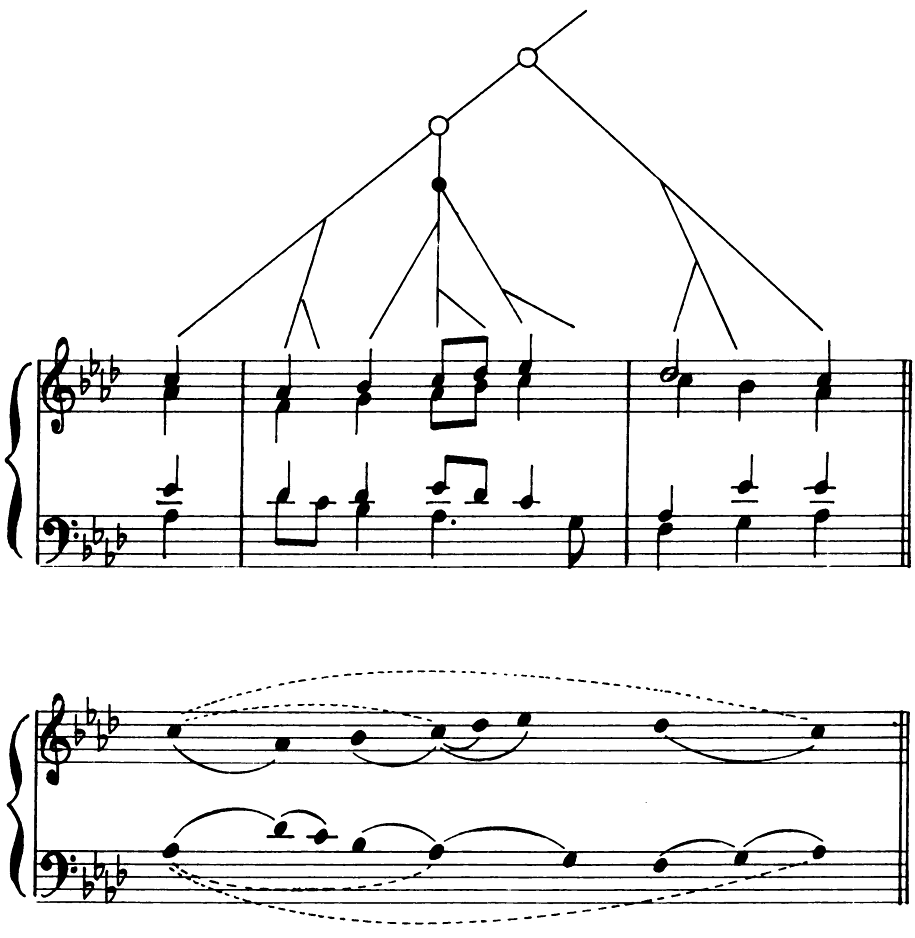


Figure 10.

gational structure (PRPR 6). In Figure 11, for instance, the third phrase of the chorale (11a) is incomplete not just because it lacks a final tonic, but because the final event is not preceded by a sequence of two events that relax into it. By contrast, the last phrase of the chorale (11b) fulfills the branching requirement for closure.[5]

Normative prolongational structure is a shaping force at the phrase level and at all larger levels of tonal architecture. The cumulative result of such

5. In the figure, the dotted lines at the top indicate that these branches attach at more global levels of the tree than is shown in the analysis (compare Figure 12). An exclusively phrasal prolongational analysis would show all branches within a phrase as attaching locally, with no branches connecting above the phrase level. We will not pursue this possibility here.

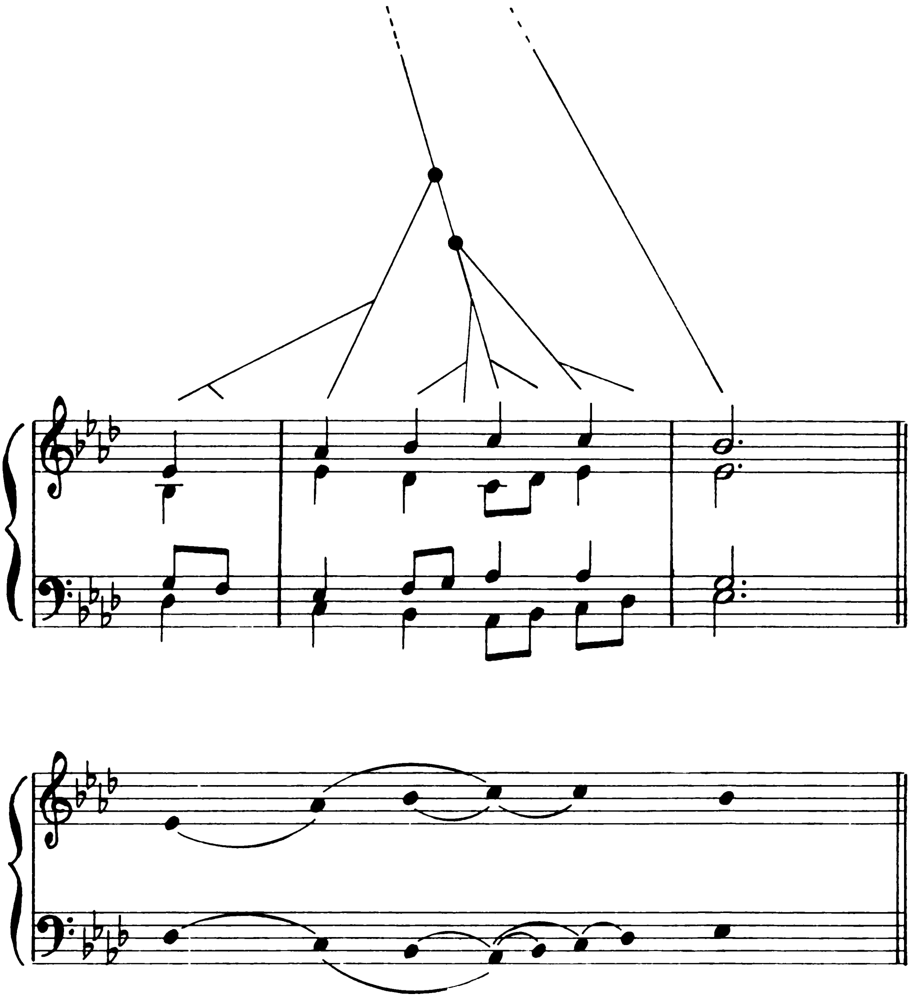


Figure 11a.

multi-leveled structures in a piece is to produce a complex tree that is at the same time filled with internal structural repetitions. The existence of such an organization appears to bear on why tonal music is rich yet learnable.

Unlike time-span trees, prolongational trees are constructed from global to local levels (i.e., from the top downwards). This is necessary because the prolongational importance of an event—its role in patterns of tension and relaxation—is determined by its larger context. Further, an event's importance cannot be evaluated solely from its pitch content; rhythmic information is also needed. In our theory, the requisite contextual and rhythmic information is not a matter of the analyst's artistic intuition, as in Schenker-

Figure 11b.

ian analysis, but is derived formally from the *time-span tree* for the piece, which in turn encodes all the rhythmic information concerning grouping and meter (PRPR 1). Thus, directly or indirectly, all four components are implicated in the description of patterns of tension and relaxation.

Besides the derivation from the time-span tree, the other critical factor in building a prolongational tree is the stability of connection between two events (PRPR 3). In a tensing pattern, the least increase in tension is the most stable connection; hence a right strong prolongation (9a) is highly favored. In a relaxing pattern, the greatest increase in relaxation is the most stable connection; hence a left progression (9f) is highly favored.

Skipping the step-by-step derivation,[6] we offer a full prolongational reduction of the chorale in Figure 12. For simplicity, the analysis proceeds only to a certain degree of detail. The musical notation of the reduction appears in two stages: graph *a* presents those connections made above the dashed line in the tree, graph *b* those made below the dashed line. As suggested by the Roman-numeral and scale-step designations at the bottom of the example, graph *a* displays the overall linear/harmonic motion of the piece. Note how the relative minor section (measures 7–10) tenses off the dominant arrival in measure 6, and how all of this is enclosed within the $\overset{\hat{3}}{\mathrm{I}}$ prolongation from measure 1 to measure 11, after which closure takes place. Graph *b* reveals more local patterns of tension and relaxation.

The reader should bear in mind that the reduction in Figure 12 is not the only "preferred" prolongational analysis. Various details are open to interpretation; and if the ambiguity in the time-span reduction discussed in connection with Figure 8 were resolved otherwise, a different (though perhaps less interesting) overall prolongational tree would result.

Discussion

We alluded above to the resemblance between our prolongational reduction and aspects of Schenkerian analysis. Though this resemblance is genuine, we should begin this section by pointing out various broad differences between Schenker's theory and ours. These remarks will lead gradually into brief discussions of the nonoverlapping condition on hierarchies and of the nonhierarchical dimensions of musical structure.

Some differences between the two theories are less a matter of conviction than of focus. For example, we have not developed a very interesting notion of voice-leading, a central concept in Schenker. The reason for this is that voice-leading as such is not hierarchical. On the other hand, our theory deals more centrally with rhythm and its relation to pitch than does Schenker's. Again, this is because grouping, meter, and time-span segmentation are hierarchical.

6. The derivation for the most part follows applications of PRPRs 1 and 3, supplemented by principles not mentioned here. Readers familiar with *GTTM* will appreciate that the strong prolongational connection of the opening tonic to the $\overset{\hat{3}}{\mathrm{I}}$ in measure 11 comes about through the *Interaction Principle*, since in the associated time-span tree this $\overset{\hat{3}}{\mathrm{I}}$ is one level below the current level of derivation. This is an excellent illustration of the Interaction Principle, supplementing those in *GTTM*.

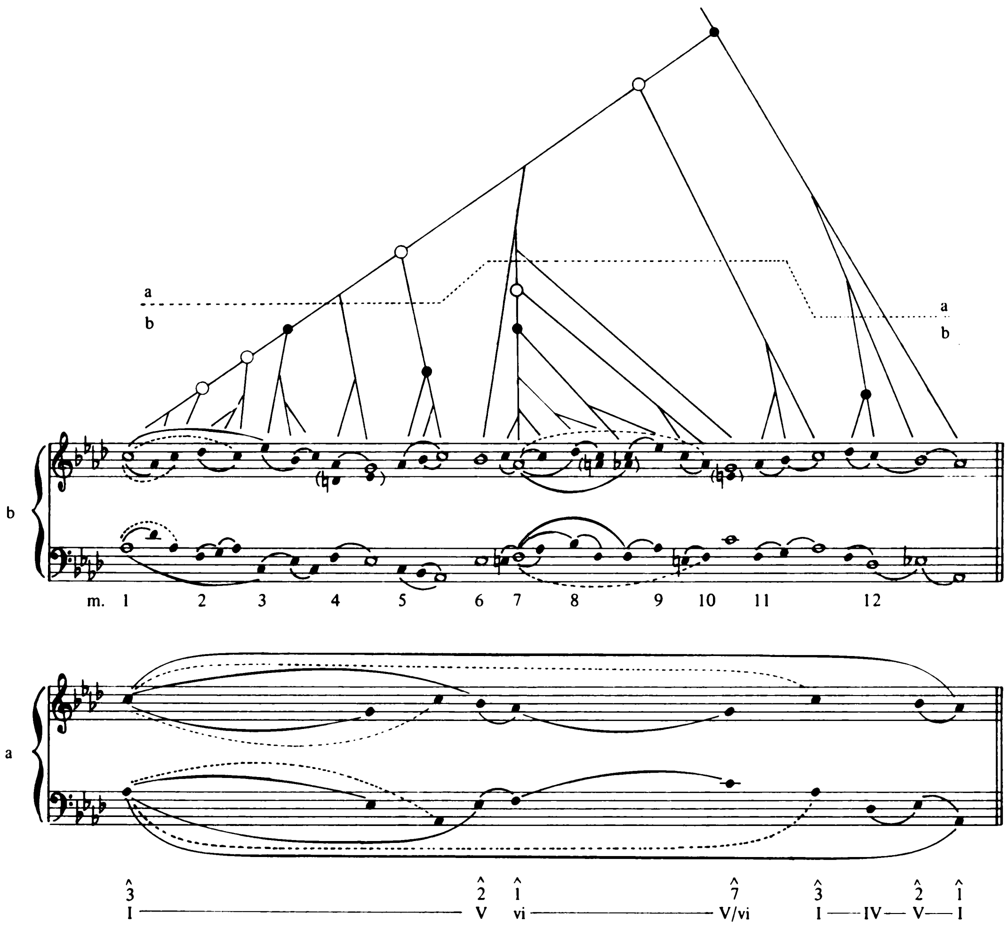


Figure 12.

A more important contrast is that Schenker's theory, though sophisticated as an analytic system, lacks any counterpart to our musical grammar (i.e., the rule system that generates structural descriptions). It is our grammar that assigns structure, regardless of what we as analysts would personally like to say about a piece. In other words, our role is not to develop analyses but to create a system that develops analyses. This approach entails (temporarily, we hope) analyses that are limited in scope and sometimes in subtlety. But it also entails the intellectually deeper gain of putting analytic principles on a formal and generalized basis.

A still more essential difference lies in the fact that Schenker's theory is aesthetic and ours is psychological. His purpose is to illuminate musical masterpieces, ours to find principles of musical cognition. All of our rules are motivated from simple, "prosaic" musical examples, of the sort a psychologist of music perception might use in an experiment. If our theory has anything to say about masterpieces, it is because listeners make sense of Beethoven by the same cognitive principles they use in understanding trivial examples. If some of these principles relate to Schenker, it is presumably because Schenker himself was at least implicitly concerned with musical cognition.

Schenker (1935) posits the basis of his theory, the *Ursatz*, as an a priori aesthetic law. Whatever the value of Narmour's (1977) criticisms of this aspect of Schenker, they do not apply to our theory, which is neither a priori nor axiomatic, but empirical. For example, the *Ursatz*-like structures at the highest branching levels of Figure 8 and Figure 12 do not result from any presumed status of these structures. Rather, they arise directly out of applications of the rules, which together predict maximally stable structures. If the piece analyzed were tonally less stable, *Ursatz*-like structures would not appear. From our vantage, Schenker's *Ursatz* simply embodies many of the stability-making features of the tonal idiom.

In a similar vein, Schenker often makes transformations on a musical surface to make it conform to his aesthetic laws. (Here we mean *transformation* in the technical sense of some operation on a sequence that changes its order or content. A hierarchical description is not a transformation.) Though we too use transformations to explicate certain musical phenomena, their role in our theory is much more restricted. An excessive use of them leads away from the data, the musical surface, and into a priori abstractions.[7]

A technical difference between the two theories concerns our strict definition of pitch hierarchy: branches must be right or left, and they cannot overlap or intersect. Though Schenker often makes one event subordinate to another, there are numerous cases where he does not. And overlapping hierarchical interpretations for different voice-leading lines are common in Schenker's looser version of reduction. Let us consider these points in turn.

It is clear that a suspension is subordinate to its resolution (as a left branch); but what about a passing or neighboring tone? One can easily imagine a hierarchical description in which some events branch and others

7. There are many examples of transformation in Schenker's (1932) analysis of "Ich bin's . . . ." One of them deserves special comment: he turns the IV^7 on the downbeat of measure 12 into a IV–I^6_4 progression at a background level. Assuming a transformation is required here, we would opt for a ii^6_5, both on stylistic grounds (Bach hardly ever employs a cadential $^6_4$ in his chorales, whereas ii^6_5–V–I is entirely typical—not that such matters would bother Schenker), and for systematic reasons (less transformation is needed, especially rhythmically, to arrive at the ii^6_5 from the surface).

are simply subordinate "insertions." We have avoided this alternative because we have been able to assign structural interpretations to right and left branching that have a psychological dimension. In time-span reduction, the branchings show the position of pitch importance within grouping and meter. In prolongational reduction, they reveal patterns of tension and relaxation (even a neighboring tone tenses or relaxes as a result of rhythmic factors). Thus the case for the branching condition seems to us strong.

The nonoverlapping condition raises a number of issues. Inherent in our tree notation is the conception of a musical surface as a sequence of discrete events. This view is only partly accurate: music is also perceived as simultaneous polyphonic lines. To describe polyphony, the grammar would have to generate multiple (but related) groupings and trees, with one structural description for each line. Each such description would nevertheless continue to observe the nonoverlapping condition. We reserve for future research an extension of the theory in this direction.

Meanwhile there are three areas in which our current theory accommodates *apparent* overlapping or intersecting structures. The first involves certain transformational rules. For example, one transformational rule analyzes grouping overlaps and elisions, permitting double branching for overlapped events; another treats functioning events that are literally missing from the musical surface (an instance appears in measure 10 of Figure 8); yet another allows the fusion of arpeggiated figures and accounts for "auditory stream segregation" in music (Bregman & Campbell, 1971). In these cases the nonoverlapping condition is weakened only in a restricted manner and only for certain well-defined phenomena.

The second area where overlapping or intersecting apparently occurs is in moments of structural ambiguity for which one might seek multiple branching interpretations. As suggested for Figure 8, however, the correct solution is not to assign crossing branches—this would tell us nothing—but to generate competing preferred structures, each of which is well-formed. The competing rule applications and descriptions thereby explicate the sense of the ambiguity.

The third area has to do with the existence of four hierarchical components in a structural description. Theorists with only one hierarchical dimension at their disposal might feel compelled to cross branches (or the equivalent) in an attempt to capture their analytic intuitions. The problem lies with their impoverished theoretical model. Once grouping and meter are conceptually disentangled and the two kinds of reduction are both seen as necessary, it becomes possible to represent a variety of musical intuitions clearly and without overlapping hierarchies. Each component remains well-formed but enters into a structural counterpoint with the others. Some flavor of this can be obtained by comparing the time-span and prolongational reductions of "Ich bin's . . ." (Figure 8 and Figure 12). If superimposed, their respective branchings would cross profusely, showing how the music both articulates (time-span reduction) and progresses (prolongational reduction):[8]

Despite these remarks, let us imagine for a moment a structure characterized by the complete absence of the nonoverlapping condition, whose only constraint is noncontradiction in domination and subordination, and in which the only criterion for relationship among its elements is some kind of quasi-elaborational connection. For illustration, consider the first, third, and last phrases of the chorale. One might observe certain similarities here, such as the close melodic connections among the three phrases, the presence of a $\overset{3}{\mathrm{I}}$ on the third beat of the first measure in each phrase, and the parallel harmonic progressions at the ends of the first and last phrases (especially the IV^7s on the downbeats of their respective second measures). Similarly, the second and fifth phrases would connect, creating a structure with crossing branches. Is such a structure possible?

In reply, we note Meyer's (1973) important distinction between *hierarchical*, *associational* (his word is "conformant"), and *implicative* structures (also see Narmour, 1977, on implicative structures). The just-mentioned connections in the chorale are not hierarchical but associational; that is, these aspects of the various phrases are perceived as related, but not in a subsuming or containing manner. A tree representation for these connections would therefore be inappropriate. (In terms of our theory, associational structures lack well-formedness conditions, and must be treated by preference rules alone.)

The chorale can also illustrate implicative structure. The melodic B♭ at the end of the third phrase implies a "realization" on A♭, which arrives at the end of the sixth phrase. The inverted IV^7–V–I progression at the end of the first phrase perhaps implies its realization in root position at the end of the piece; the effect of this relationship, in any case, is powerful.

Other details aside, it is apparent that associations and implications are strongly influenced by hierarchical function. For example, the final B♭ in the third phrase resolves to the final A♭ in the last phrase, not to any of the intervening A♭s, because of their respective importance and function in the time-span reduction. This suggests that an implicative theory must take as its input not just the musical surface but the associated time-span reduction. Conversely, our rules for parallelism in all four components demand greater specificity from an adequate associational theory.

In developing our theory of musical cognition, we have found it method-

8. Incidentally, multiple tree descriptions are also necessary in linguistic theory. Recent research has shown that sentences are assigned both a syntactic tree and a phonological tree (see Liberman & Prince, 1977; and *GTTM*).

ologically fruitful to concentrate on hierarchical structures at the expense of associational and implicative structures. The well-formedness conditions impose the kind of constraints in theory-building that enables progress to be made. With this relatively firm foundation, it may be possible to approach the other, more open-ended dimensions in a rule-governed way.

References

Bregman, A. A., & Campbell, J. Primary auditory stream segregation and perception of order in rapid sequences of tone. *Journal of Experimental Psychology*, 1971, ***89***, 244–249.

Chomsky, N. *Aspects of the theory of syntax*. Cambridge, Mass.: M.I.T. Press, 1965.

Cooper, G., & Meyer, L. B. *The rhythmic structure of music*. University of Chicago Press, 1960.

Deutsch, D., & Feroe, J. The internal representation of pitch sequences in tonal music. *Psychological Review*, 1981, ***88***, 503–522.

Jackendoff, R. *Semantics and cognition*. Cambridge, Mass.: M.I.T. Press, in press.

Jackendoff, R., & Lerdahl, F. Generative music theory and its relation to psychology. *Journal of Music Theory*, 1981, ***25***, 45–90.

Jackendoff, R., & Lerdahl, F. A grammatical parallel between music and language. In M. Clynes (Ed.), *Music, mind, and brain*. New York: Plenum Press, 1982.

Keiler, A. The syntax of prolongation I. *In Theory Only*, 1977, ***3.5***, 3–27.

Koffka, K. *Principles of Gestalt psychology*. New York: Harcourt, Brace & World, 1935.

Lerdahl, F., & Jackendoff, R. Toward a formal theory of tonal music. *Journal of Music Theory*, 1977, ***21***, 111–171.

Lerdahl, F., & Jackendoff, R. On the theory of grouping and meter. *The Musical Quarterly*, 1981, ***67***, 479–506.

Lerdahl, F., & Jackendoff, R. *A generative theory of tonal music*. Cambridge, Mass.: M.I.T. Press, 1983.

Liberman, M., & Prince, A. On stress and linguistic rhythm. *Linguistic Inquiry*, 1977, ***8***, 249–336.

Marr, D. *Vision*. San Francisco: W. H. Freeman & Co., 1981.

Meyer, L. B. *Explaining music*. Berkeley: University of California Press, 1973.

Narmour, E. *Beyond Schenkerism*. University of Chicago Press, 1977.

Schenker, H. *Fünf Urlinie-Tafeln*. New York: Dover Publications, Inc., 1969. (Originally published, Vienna: Universal Edition, 1932.)

Schenker, H. [*Der freie Satz*] (E. Oster, Ed. and trans.) New York: Longman, Inc., 1979. (Originally published, Vienna: Universal Edition, 1935.)

Wittgenstein, L. *Philosophical investigations*. Oxford: Blackwell, 1953.