Composition, Perception, and Schenkerian Theory

DAVID TEMPERLEY

In this essay I consider how Schenkerian theory might be evaluated as a theory of composition (describing composers' mental representations) and as a theory of perception (describing listeners' mental representations). I propose to evaluate the theory in the usual way: by examining its predictions and seeing if they are true. The first problem is simply to interpret and formulate the theory in such a way that substantive, testable predictions can be made. While I consider some empirical evidence that bears on these predictions, my approach is, for the most part, informal and intuitive: I simply present my own thoughts as to which of the theory's possible predictions seem most promising—that is, which ones seem from informal observation to be borne out in ways that support the theory.

Keywords: Schenkerian theory, music perception, music composition, harmony, counterpoint

I. THE QUESTION OF PURPOSE*

More than seventy years after Schenker's death, Schenkerian theory remains the dominant approach to the analysis of tonal music in the English-speaking world. Schenkerian analyses seem as plentiful as ever in music theory journals and conference presentations. Recent bibliographical works,[1] translations of Schenker's writings,[2] and extensions of the theory to non-canonical repertories[3] testify to the theory's continuing vitality. In music theory pedagogy, Schenker remains a powerful and pervasive force. Several recent textbooks either focus on the theory directly[4] or present basic harmony and counterpoint from a Schenkerian perspective.[5] Schenkerian analysis remains the standard next step in tonal analysis after the rudiments of the undergraduate core; at many institutions, including my own, graduate or upper-level undergraduate courses on "tonal analysis" are exclusively devoted to Schenkerian analysis.

No theory should be exempt from critical scrutiny; such scrutiny is especially important if a theory is highly influential and widely accepted. Schenkerian theory (hereafter ST) has certainly not escaped criticism over the years. Numerous critiques of the theory have been offered,[6] to which its advocates have responded in vigorous defense.[7] I will argue here, however, that some basic questions about ST have not yet been asked—or, at least, have not been given the careful and sustained attention that they deserve.

* Thanks to Nicholas Temperley for helpful comments on an earlier version of this essay.

1 E.g., Ayotte (2004); Berry (2004).

2 Schenker ([1924] 2004); ([1930] 1994).

3 E.g., Judd (1992); Covach and Boone (1997); Larson (1998); and Maisel (1999).

4 Cadwallader and Gagné (2007).

5 Roig-Francoli (2003); Gauldin (2004).

6 Rosen (1971); Narmour (1977); Smith (1986).

7 Schachter (1976); Keiler (1978); Beach (1987).

A theory is (normally, at least) designed to make accurate predictions about something; we evaluate it by examining how well its predictions are confirmed. Thus, the first step toward evaluating ST is to determine what exactly its predictions are about—that is, what it is a theory of. This is, in fact, an extremely difficult and controversial issue. In some respects, there is general agreement about the basic tenets of ST. By all accounts, the theory posits a hierarchical structure in which musical events are elaborated by other events in a recursive fashion: Event X is elaborated by event Y, which in turn may be elaborated by event Z. (This process of elaboration is sometimes also called prolongation, diminution, or transformation.) The elaborating and elaborated events may be pitches, although in some cases they may also be more abstract entities, notably structural harmonies or *Stufen*. Viewed generatively, the process begins with the *Ursatz*—a descending stepwise melody $\hat{3}$–$\hat{1}$, $\hat{5}$–$\hat{1}$, or $\hat{8}$–$\hat{1}$, over a I–V–I harmonic progression—and ends with the pitch structure of the actual piece. A Schenkerian analysis thus forms a kind of "tree" structure, similar to the syntactic trees posited in theoretical linguistics. There is widespread agreement, also, that ST is primarily a theory of common-practice tonal music—certainly this was Schenker's intention—though some have argued that it is applicable to other kinds of music as well, such as pre-tonal Western music, post-tonal music, jazz, and rock.

ST, then, posits a certain kind of musical structure—hereafter I will call these Schenkerian structures. The difficulty lies in determining what, or perhaps more precisely where, these structures purport to be—how they map onto the real world. To repeat my earlier question: What exactly is ST a theory of?[8] Two general possibilities are that it could be a theory of composition or a theory of perception. As a theory of composition, the theory states that Schenkerian structures are mental representations formed in the minds of composers as part of the process of creating a piece. This view does not necessarily imply that pieces are

8 For a more thorough discussion of this issue—which I treat only very briefly here—see Temperley (1999).

generated in a top-down order—starting with the *Ursatz* and proceeding downward—but simply that the entire Schenkerian structure is generated at some point during the compositional process. The compositional view of ST also does not imply that composers were necessarily consciously aware of Schenkerian structures; it is a basic tenet of cognitive science that many mental processes and representations are not consciously accessible. Schenker himself clearly advocated ST as a theory of the compositional process (though he believed that only great composers of tonal music, not mediocre ones, had Schenkerian structures in mind).[9] Today, one rarely finds explicit endorsements of ST as a theory of composition, but it is certainly implicit in much Schenkerian writing. For example, a statement such as "Beethoven recomposes the *untransposed* falling sixth . . . over the span of 13 bars in the development section"[10] surely constitutes a claim about the compositional process. And some authors, notably Brown,[11] have claimed quite forthrightly that ST should be construed as a theory of expert tonal composition.

The second possibility is that ST could be construed as a theory of perception. By this view, in hearing a piece, listeners infer a Schenkerian structure for it. As with the compositional view, the perceptual view of ST does not necessarily imply that Schenkerian structures are available to consciousness; they may well be present at an unconscious level. The perceptual view of ST, like the compositional view, is rarely explicitly embraced; indeed, some authors have argued against it.[12] But again, one often finds it implied in Schenkerian writings, and I suspect it is actually quite widely held (which is probably why some people bother to argue against it!). In particular, as I have argued elsewhere,[13] claims that Schenkerian structures *explain* certain aspects of our perception and experience—the sense of coherence we may experience in a piece, for example—seem to imply that such structures are being mentally represented.

The compositional and perceptual views of ST are not mutually exclusive. A parallel here comes to mind with linguistics (one of many that will be drawn in this essay). When linguists posit syntactic tree structures, for example, the usual assumption is that such structures play a role in both production and comprehension: speakers form them in creating a sentence, and listeners form them in understanding the sentence. Similarly, one might argue that Schenkerian structures characterize both composition and perception. In support of this view, one might argue that the purpose of music, after all, is communication. Why would the great composers have bothered to create such elaborate mental structures if they thought that these structures would never be shared by listeners? Thus, the possibility of construing ST as a theory of both composition and perception should be borne in mind.

The two possible views of ST presented above—as a theory of perception or a theory of composition (or both)—are not the only ways in which the theory might be construed. Another possibility is that a Schenkerian analysis could simply be regarded as an interesting or satisfying way of hearing a piece, not entailing any claim either about listeners' existing perceptions or about the compositional process. Indeed, I would argue that this view—which I have elsewhere called the "suggestive" view—is in fact the dominant construal of the theory in contemporary Schenkerian discourse, or at least the one that is most often explicitly advocated.[14] This construal of ST is perfectly coherent, and the goal it envisions for the theory is a very worthwhile one. Indeed, it has a direct benefit that the perceptual and compositional goals do not: it enhances our enjoyment and appreciation of a piece of music. But this is a very different goal from the compositional or perceptual construals of the theory; indeed, one might question whether, under the suggestive construal, ST is a "theory" at all. Construed suggestively, the aim of the theory is not to generate true statements or correct predictions about anything, but simply to give rise to a satisfying experience for the listener. Perhaps the theory could, in some way, be systematically evaluated as to how well it achieved the goal, but that is not my intent here. For now, I will give no further attention to the suggestive view of ST and will focus on the perceptual and compositional construals.[15]

In what follows I will consider—in a very preliminary and speculative way—how Schenkerian theory might be tested both as a theory of composition and a theory of perception. I propose to evaluate the theory as we would evaluate any theory: by examining its predictions and seeing if they are true. The first problem—and really the primary concern of this essay—is simply to interpret and formulate the theory in such a way that substantive, testable predictions can be made. I will consider some evidence from corpus and experimental research that seems to bear on these predictions (at present, little such evidence is available). But for the most part, my approach in this essay will be quite informal and intuitive: I simply present my own thoughts as to which of the theory's possible predictions seem most promising—that is, which ones seem from informal observation to be borne out in ways that support the theory.

9 See, for example, Schenker's remarks in *Free Composition* ([1935] 1979, xxii, xxiii, 6, 128–29).

10 Cadwallader and Gagné (2007, 7).

11 Brown (2005, 222–33).

12 See for example Schachter (1976, 285–86) and (1981, 122–23); and Benjamin (1981, 160, 165).

13 Temperley (1999, 70–71).

14 See Ibid. (69–75) for discussion of the "suggestive" view of analysis and examples from the Schenkerian literature.

15 One other possibility is that ST could be regarded as an objective theory of musical structure, without entailing any claim about either perception (either descriptively or suggestively) or composition. This brings to mind the semiotic approach of Nattiez (1990, 11–12), which distinguishes between the poietic level of musical structure (in the mind of the producer), the esthesic level (in the mind of the perceiver), and the neutral level, which is simply the objective structure of the piece itself: the latter would seem to correspond to the goal that I have just articulated. I do not believe that this view of ST is widespread in current thinking, though it is difficult to be sure.

There is certainly more to be said about the relationships—conflicts and convergences—between these various goals. For example, the suggestive and compositional goals of analysis may often converge: the composer's conception of a piece is likely to be an interesting and satisfying way of hearing it, though perhaps not always the most satisfying way.

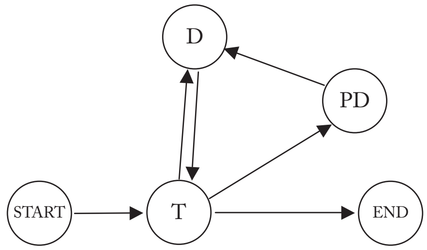
Such speculative reasoning is of course no substitute for more rigorous methods, and I hope that this study will lay the groundwork for such testing in the future. But intuitive hunches can often be of value in the development and testing of theories, as they may guide us toward the most promising avenues for further exploration.

2. schenkerian theory as a theory of compositional practice

2.1 predictions about musical corpora

In this section I consider how Schenkerian theory might be evaluated as a theory of composers' mental representations. One might argue that testing a theory in this way is not just difficult but impossible. The creative processes that gave rise to common-practice pieces are long gone, as are the composers themselves, and no longer available for study. However, we can still study the compositional objects themselves—actual pieces of music—as evidence of the compositional process. In fact, I would argue that this reasoning is quite commonplace. As I suggested earlier, music-theoretical and analytical discourse often contains implicit claims about the compositional process. Most often, these claims are not based on direct historical evidence, but simply on musical evidence: they are justified by the music itself. In other words, music itself can be taken as a kind of evidence about compositional thought; the best model of the compositional processes involved in creating a corpus of music is the one that makes the most accurate predictions about that corpus.

To illustrate this approach, let us consider a simple example from the realm of harmony. A basic principle of common-practice harmony is that harmonies belong to functional categories—tonic, dominant, and predominant—and the arrangement of these categories tends to follow certain patterns: dominants (V and VII) go to tonics (I); predominants (II and IV) go to dominants; and tonics can go to any category. (III and VI are functionally problematic, and I will disregard them here.) These patterns can be expressed in what is known as a finite-state model, as shown in Example 1. The model consists of a set of nodes representing "states," connected with lines representing "transitions"; it begins in the "start" state and then moves from state to state in any way that the transitions will allow. It can be seen that the progression T–D–T is generated by the model, as are T–PD–D–T and T–D–T–PD–D–T; other progressions, such as PD–T–D and T–D–PD–T, are not generated. We can use this model to represent a prediction about common-practice harmony, namely, that only progressions that can be generated by the model will occur.



example 1. *A finite-state model of functional harmony (T=tonic, D=dominant, PD=predominant)*

An immediate problem with this model is that its predictions do not always hold: sometimes predominants move to tonics, for example, as in a plagal cadence. Admittedly, such exceptions show that functional harmonic theory is imperfect as a model of tonal harmony; but they do not show that it is useless. A theory whose predictions hold true most of the time can still be of great value; we use such theories all the time in our daily lives. Imperfect though it may be, functional harmonic theory represents a powerful and valid generalization about tonal harmony, better than many conceivable alternatives—for example, a theory that posited that chords are chosen at random without regard for the previous chord, or that predominants move to tonics and dominants move to predominants. On this basis, I would argue, we are justified in positing functional harmony as part of the knowledge that common-practice composers brought to bear in their compositional process.

I have suggested that the predictions of harmonic theory are true "most of the time." In a study that is all about the rigorous testing of theories, it would hardly be fair for me to make such a claim without some kind of empirical verification. A small amount of relevant data is provided in Example 2. The table shows counts of two-chord harmonic progressions in a corpus of forty-six common-practice excerpts from the workbook to Stefan Kostka and Dorothy Payne's textbook *Tonal Harmony* (1995); an instructor's manual[16] provides harmonic analyses by the textbook authors. While the data do not adhere completely to traditional principles of harmonic theory, there is clearly a close correspondence. The most frequent progressions in the data (in order) are V–I, I–V, II–V, I–IV, and I–II, all normative progressions of harmonic theory; non-normative progressions, such as II–I and V–IV, are generally much less common. Thus this data seems to give considerable support to the predictions of harmonic theory.[17] While the choice of harmonic progressions clearly was not dictated *entirely* by the principles represented in Example 1, the data suggest that these principles did play some role in the creative process.

Basic tonal theory abounds with generalizations about compositional practice. For example, parallel perfect intervals tend to be avoided; chordal sevenths resolve downward by step; the largest interval in a chord tends to be at the bottom; phrases tend to be four or eight measures long; and so forth. As with the

16 Kostka (1995).

17 For more about the Kostka-Payne corpus and the way this test was conducted, see Temperley (2009). One problem with the test is that the excerpts in the Kostka-Payne workbook may have been selected by the authors, in part, to support a particular theoretical viewpoint, and therefore may be a somewhat biased sample. This is an important point to bear in mind with corpus analysis in general.

	Cons											
Ant	I	♭II	II	♭III	III	IV	♯IV	V	♭VI	VI	♭VII	VII
I	0	7	31	1	4	45	2	84	11	17	3	19
♭II	2	0	8	0	0	0	1	3	0	0	0	1
II	5	3	0	1	4	1	7	62	2	8	0	6
♭III	1	1	0	0	0	0	0	4	4	0	0	0
III	1	0	2	0	0	7	0	1	0	7	0	1
IV	27	2	10	0	4	0	3	16	0	1	1	4
♯IV	3	0	0	0	0	0	0	13	0	0	0	0
V	166	0	8	1	2	4	0	0	7	6	0	2
♭VI	3	2	8	0	1	3	0	4	0	3	2	0
VI	4	2	28	0	1	4	2	1	0	0	0	1
♭VII	0	0	0	5	0	0	0	1	0	0	0	0
VII	26	0	0	0	3	0	1	2	1	0	0	0

EXAMPLE 2. *Harmonic progressions in the Kostka-Payne corpus*
The data represent both major-key and minor-key sections. Antecedent chords are on the vertical axis, consequent chords on the horizontal; each cell in the table shows the number of occurrences of the antecedent-consequent pair. All chords were categorized as labeled in the Roman numeral analyses (i.e., in relation to the local key), with the assumption of enharmonic equivalence (e.g., ♯I = ♭II); applied chords were relabeled as immediate chords of the local key (e.g., V/V as II). Cadential six-four chords are notated in the analyses in a two-level fashion, with a "I^6_4–V" within a larger "V"; the latter analysis was used here.

principles of functional harmonic theory, these principles are not held to be inviolable laws, but, rather, generalizations that are mostly true; and on that basis, they are—quite legitimately—widely accepted to have been part of compositional thinking. Of course, such claims are generally based not on systematic data analysis, but on informal observation—passed on and reconfirmed over many generations. In recent years, however, a number of these principles have been subjected to empirical testing, generally with positive results.[18]

Once appropriate data have been gathered, the question arises: how can we judge whether, or to what degree, the data confirm the theory's predictions? One can of course judge this intuitively—as I did in saying that the data in Example 2 give "considerable support" to harmonic theory—but clearly such statements are vague and informal; one might wonder if the "fit" between theory and data could be measured more objectively. There is in fact an extremely powerful method for doing this, namely, Bayesian probabilistic modeling. Under the probabilistic approach, the validity of a theory of a body of data depends on the probability that it assigns to the data. A finite-state model, for example, can be made probabilistic by assigning probabilities to the transitions; the probability of a series of output symbols is then the product of the probabilities of all the transitions. By this reasoning, the best model of the data—and thus the one most likely to characterize the process of their creation—is the one that assigns the data highest probability. Elsewhere I have explored the possibility of using this approach to evaluate claims about compositional processes.[19] Conceivably, this approach could be applied to ST as well. I do not attempt this here, and I believe it would be premature to do so, for reasons that will become apparent below. But the probabilistic approach is still worth bearing in mind, as it suggests a way in which even informal and qualitative empirical statements—e.g., "dominants usually go to tonics"—might ultimately be subjected to quantification and rigorous testing.

To summarize the previous discussion: A theory can be regarded as a valid theory of compositional practice to the extent that it makes true predictions—not necessarily perfectly true, but mostly true—about compositions. We now apply this approach to ST. What exactly are the predictions of ST? And are they true?

2.2 PRELIMINARY ISSUES

Before proceeding, we must consider several basic issues in the testing of ST. One is the issue of musical domain. If ST is a theory of the compositional process, whose compositional process does it purport to describe? While there is some diversity of opinion on this point, the usual view seems to be that ST is a theory of tonal music[20] or some more narrowly defined subcategory such as "functional monotonality"[21] or "triadic tonality."[22] As a theory of compositional processes, then, it presumably applies to composers of such music. Those interested in testing ST

18 The work of David Huron is especially notable in this regard. See, for example, Huron (2001) with regard to chord spacing (17–18) and the avoidance of perfect parallel intervals (31).

19 Temperley (2007, 159–79); Temperley (2010). See also Conklin and Witten (1995); and Mavromatis (2005).

20 Cadwallader and Gagné (2007, xi).

21 Brown (2005, 22, 24).

22 Forte (1977, 25).

have recognized the importance of defining the domain to which the theory applies. In Matthew Brown's words:

> [S]ince laws and theories delimit classes of phenomena, they must have boundaries. . . . [A] valid theory of tonality must be able to predict specific pitch relationships that occur in tonal pieces and not in non-tonal ones.[23]

Here Brown voices a valid and important concern. A theory must be judged not only on its ability to generate all the observed data, but on its ability *not* to generate other possible (but not observed) data. To return to our harmonic finite-state model: If we added transitions from every node in the model to every other node (for example, connecting D to PD and connecting PD to T), the model would then generate every possible harmonic progression. It would still of course generate all of the progressions that were generated before, but it would now allow many progressions that are uncharacteristic of common-practice harmony (or, at least, rare). Such a model clearly does not predict the data well, and therefore has little plausibility as a model of the compositional process. The crucial question to ask of ST, then, is not so much whether it can generate the observed data, but whether it can discriminate between what is observed and what is not.[24]

How, then, should the domain of ST be defined? Defining this domain as "tonal" music, or something similar, is certainly a possibility; this may not be the best approach, however. First of all, whether a piece is "tonal" is highly subjective and, I would suggest, quite ambiguous. Surely we cannot draw a sharp line between music that is in this category and music that is not.[25] Secondly, it is not crucial to define the domain of the theory in the largest possible way. Suppose, for example, that we confined ourselves at least initially to music written in Europe between 1770 and 1800. This is surely a well-defined domain; there is no doubt (barring a few pieces of uncertain date or composer) as to which pieces it contains. It certainly does not contain everything to which ST would normally be considered applicable, but it is a start. If one had a theory about dogs, and it later turned out that the theory was true of all mammals, this would not invalidate the theory as a theory of dogs (though it would mean that the domain of the theory could be more accurately characterized). One might argue, indeed, that it is more prudent to begin by trying to model a relatively small and unproblematic domain rather than a large and ambitious one.

There is no need to resolve the question of domain definitively here, given the preliminary nature of this investigation. My aim has only been to consider some of the issues that arise. Rigorous, quantitative testing of ST will require resolution of these issues, but this is a project for the future. For now, I will follow convention and simply refer to the domain of ST as "tonal music."

Another issue to consider is the terms in which the theory's predictions will be expressed. The important point here is that, for a theory to be convincing, it must make predictions about entities that are themselves not in question—that is to say, it must predict accepted facts. To use the language of the philosophy of science, the theory's predictions must concern "observation terms"—things that are directly observable and hence uncontroversial.[26] Return again to the finite-state model in Example 1, which makes predictions about harmonic progressions. Such a model presupposes that the sequence of chords in a piece is a matter of fact, about which all qualified observers will agree. Clearly this assumption is not completely true—there is sometimes disagreement as to how to assign Roman numerals to a piece (for example, what is a chord and what is merely an ornamental or linear event). This is a potential problem for harmonic theory as a model of compositional practice. To address this problem, the theory would have to be supplemented with additional principles that relate harmonies to the "observable facts" of music: first and foremost, actual patterns of notes. And this points to an important virtue of ST: clearly, the theory *does* address patterns of notes, and the relationship between note patterns and higher-level structures, in a way that functional harmonic theory does not. Any formalization of ST as a predictive theory should surely take advantage of this aspect of the theory. By contrast, a model of ST that makes predictions only about entities posited by the theory itself will have little persuasive force. For example, the prediction that "every piece is built on either a $\hat{3}$–$\hat{1}$, $\hat{5}$–$\hat{1}$, or $\hat{8}$–$\hat{1}$ *Ursatz*" will cut little ice with anyone who is skeptical about the theory, since they will question the very "facts" the theory is predicting.

This issue of observation terms arises in a study by Brown, Douglas Dempster, and Dave Headlam.[27] These authors suggest that ST makes a prediction regarding *Stufen* (structural harmonies), namely, that a I *Stufe* is never directly elaborated by a ♯IV/♭V. For those who take for granted the presence of *Stufen* and the relations of elaboration between them, this is indeed a testable prediction. For someone who is skeptical about ST,

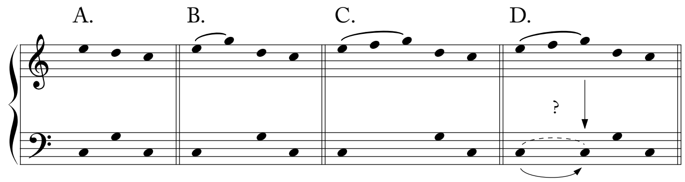
23 Brown (1989, 29).

24 Probabilistic methods offer an elegant solution to this problem. As noted earlier, a probabilistic model assigns a probability to a body of data, and models can be evaluated and compared using these probabilities: the higher the probability, the better the model. By definition, the probabilities assigned to all possible events by a probabilistic model (where events might be pieces, harmonic progressions, or anything else) must sum to 1. Thus if a model assigns relatively high probability to events that occur (or occur frequently) in a musical corpus, it must assign relatively low probability to other events. For this reason, there is no need to actually test a model on its ability to assign low probability to events not occurring in the corpus; one need only test it on the corpus itself.

25 In effect, defining the domain as "tonal music" appeals to our intuitions about what is tonal and what is not. But the basis for these intuitions is an interesting perceptual problem in itself; what is it that makes a piece sound or seem "tonal"? One could certainly devise a model that predicted these judgments; this would be an interesting project. (Indeed, ST itself might be used for this purpose, as I will discuss.) But I would argue that modeling compositional data, and predicting judgments of tonality, are two different problems and should be kept separate. The domain of a compositional theory might roughly correspond to music that seems tonal, but to *define* the domain in those terms seems ill-advised.

26 Fodor (1984, 25, 38).

27 Brown, Dempster, and Headlam (1997).



EXAMPLE 3. *Generating a Schenkerian reduction with a context-free grammar*

however, the reality of *Stufen* is exactly what is at issue. To convince such a skeptic, it would need to be shown how the ♯IV/♭V hypothesis led to predictions about actual note patterns—or, perhaps, about surface harmonies or other entities whose validity was not in question.

2.3 SCHENKERIAN THEORY AS A CONTEXT-FREE GRAMMAR

As stated at the end of Section 2.1, my aim is to formulate Schenkerian theory in a way that yields testable predictions about what occurs and does not occur in tonal music. I begin by considering the idea of formulating the theory as a context-free grammar, a kind of hierarchical model widely used in linguistics. After reaching skeptical conclusions about this approach, I consider some other approaches that might be used in testing the compositional validity of ST.

As mentioned above, an essential part of ST is its hierarchical, recursive nature, and it seems appropriate to try to formulate the theory in a way that reflects this aspect. Let us suppose we begin with the *Ursatz*—a $\hat{3}$–$\hat{2}$–$\hat{1}$ linear progression over a $\hat{1}$–$\hat{5}$–$\hat{1}$ bass line (other forms of the *Ursatz* will not be considered here). Each element of the *Ursatz* can be expanded or transformed to introduce new elements; then these elements in turn can be elaborated (see Example 3). For example, the initial $\hat{3}$ of the melody might expand into a $\hat{3}$–$\hat{5}$ arpeggiation, as shown in Example 3(b), or into a neighbor motion $\hat{3}$–$\hat{4}$–$\hat{3}$. (The possibility of a single event undergoing multiple elaborations would also have to be considered.) In some cases, an elaboration might be generated from two superordinate events; passing-tones and other linear progressions—where a consonant interval is filled in with stepwise motion—could be generated in this way, as shown in Example 3(c).[28] There might also be elaborations of a vertical nature: for example, building a triad under an existing melody note or over a bass note (the Schenkerian idea of "consonant support").

What I have just described is a context-free grammar (CFG), a widely used formalism in the modeling of natural language.[29] The grammar is "context-free" because the possible expansions of an element depend only on the element itself, and not on any larger context. The construal of ST as a context-free grammar has antecedents in important attempts to formalize the theory by Michael Kassler and Brown.[30] Kassler proposes a model that generates Schenkerian structures through the recursive application of vertical and horizontal elaborations, and shows how a tonal piece—or more specifically a fairly foreground-level Schenkerian analysis of that piece—can be generated in this manner. Brown, similarly, formulates ST as a set of transformational rules which can be recursively applied to an *Ursatz* to generate tonal surfaces. Both Kassler and Brown state as their goal the development of a grammar that can generate all and only tonal pieces.

In considering how to formalize Schenkerian theory as a CFG, we immediately encounter a number of difficulties. One concerns the interaction between the vertical and the horizontal. In Example 3, suppose one generated a $\hat{3}$–$\hat{4}$–$\hat{5}$ pattern from the initial E_5, and then wished to give "consonant support" to the $\hat{5}$ by adding a C_3 in the bass, as shown in Example 3(d). Would this be generated horizontally (from the previous C_3), vertically (from the simultaneous G_5), or both? Probably both, as the generation of such notes seems to depend on both vertical and horizontal connections. This is a major complication that cannot easily be addressed within the framework of context-free grammars.[31] This is, however, not a problem that is unique to Schenkerian analysis; even under a more conventional view of harmony and counterpoint, the placement of notes is conditional on both horizontal and vertical considerations, and these complex dependencies would certainly have to be addressed in some way. Another issue is how the generation of *Stufen* should be integrated with the generation of notes. One possibility—adopted by Kassler[32]—is that *Stufen* could be viewed simply as emergent features of the note-generation process: for example, the bass note of a *Stufe* could first be generated horizontally, and the upper notes of the chord generated from the bass note. In that case, no explicit recognition of *Stufen* would be needed in the generative process. This does not seem very satisfactory, however; one might argue that, once a *Stufe* is

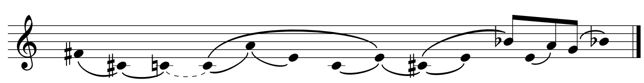
28 Allowing an event to be generated from two superordinate events might be considered to violate the requirement of "context-freeness" (as defined below). However, this does not seem like a serious problem, as the generating context in this case is still relatively local. Alternatively, one could begin with a single pitch, forming one endpoint of the progression, and then generate the other endpoint and the internal pitches with a single rule—for example, generating Examples 3(b) and 3(c) in one step.

29 This term is usually attributed to Chomsky (1957), though he used the term "phrase structure grammar."

30 Kassler (1977); Brown (1989) and (2005). Two other interesting computational implementations of ST, which came to my attention too late to be discussed here, are Gilbert and Conklin (2007) and Marsden (2010).

31 Brown (2005, 216–17) discusses other phenomena of tonal music that cannot easily be modeled within a context-free grammar.

32 Kassler (1977).



example 4. *A random fifteen-note melody generated from an even distribution of the notes C_4–B_4 (from Temperley [2007]): The analysis shown is one of 324,977 generated by a Schenkerian "parser" which finds all possible analyses using just one-note elaborations (incomplete neighbors and consonant skips). The beamed notes indicate the highest-level line or* Urlinie*; elaborations (shown with slurs) can then be traced recursively from these notes (for example, the first B♭ is elaborated by the C♯ two notes earlier, which is in turn elaborated by the Es on either side).*

generated, subsequent elaborations (such as linear progressions) depend on the *Stufe* itself rather than on its individual pitches.

Another issue that arises here (and in the modeling of ST generally) is the role of counterpoint. Schenker and his followers maintain that strict counterpoint—for example, as codified by Fux[33]—has a profound connection to common-practice composition. From the point of view of modeling composition, this is an interesting connection, since the rules of strict counterpoint are quite well defined and restrictive; it is natural to wonder whether these rules might be incorporated into a generative Schenkerian "grammar." However, the exact nature of the relationship between strict counterpoint and ST is unclear. Some, such as John Peel and Wayne Slawson, have gone so far as to claim that the higher levels of a Schenkerian analysis are strictly governed by traditional contrapuntal rules.[34] Most other authors have argued for a more indirect relationship between counterpoint and "free composition"—observing, for example, that both idioms are constructed from fluent, largely stepwise lines, and that they share certain basic patterns such as passing tones.[35]

Clearly, the formalization of ST as a CFG would entail a number of difficult challenges, but this should hardly surprise us, and is itself not a decisive argument against the approach. A further problem that arises is, in my view, more fundamental. The problem is that a Schenkerian CFG seems in serious danger of being able to generate everything. That is, it seems that almost any possible pitch pattern could be generated using some combination of basic Schenkerian transformations. In a recent experiment of my own, I showed that a completely random pitch pattern could be generated using a fairly small set of Schenkerian transformations—just single-note elaborations (incomplete neighbors) or consonant skips (see Example 4).[36] Admittedly the generative model assumed in this experiment is very far removed from actual ST, mainly because it only allows monophonic pitch sequences, thus completely ignoring the vertical dimension that (as noted earlier) is a crucial part of the theory.[37] But the essential point remains: Even if polyphonic structures were assumed, it is unclear how a Schenkerian CFG would impose substantive constraints on what surface note patterns could be generated. One possible way to address this problem would be to add probabilities to the transformations; by the usual logic of probability, the probability of a note pattern would be given by the sum of the probabilities of all of its derivations.[38] A piece would then be judged as "ungrammatical" if there was no way of generating it that was high in probability. The prospects of success for such a project are uncertain; it is unclear what kinds of regularities the grammar would be expected to capture.

This leads us to what is really the most fundamental question of all: *What is it about tonal music that a Schenkerian CFG is supposed to predict?* In what way will the model make substantive, testable predictions about tonal music, over and above what could be done with more conventional principles of harmony and voice leading? To answer this question, it is perhaps useful to consider what the motivation is for using CFGs with natural languages. This issue has been much discussed in linguistics. Noam Chomsky famously asserted that natural languages must involve something more powerful than finite-state machines, because they permit infinite recursions that cannot possibly be captured by a finite-state machine—sentences such as the following: *The man who said that [the man who said that [the man who said that [. . .] is arriving today] is arriving today] is arriving today*.[39] Sentences such as this (and other similar constructions problematic for finite-state models) involve "long-distance dependencies"—in which an element seems to license or require some element much later in the sentence. If there were analogous cases in music, this might warrant the kind of context-free grammar suggested by Schenkerian theory. Do such long-distance dependencies arise in music? Perhaps the first example that comes to mind is the *Ursatz* itself. The *Ursatz* reflects the principle that a tonal piece or tonally closed section typically begins with a I chord and ends with V–I, though the initial and final events may be separated by an arbitrarily long span of music. However, this dependency might also be captured in

33 Fux ([1725] 1971).
34 Peel and Slawson (1984, 278, 287).
35 Forte and Gilbert (1980, 41–49); Cadwallader and Gagné (2007, 23–39).
36 Temperley (2007, 176).
37 Schenkerian theory is really polyphonic in its very essence; even monophonic pieces such as the Bach violin and cello suites are generally understood in Schenkerian analysis to have an implied polyphonic structure.
38 Probabilistic context-free grammars are widely used in computational linguistics (Manning and Schütze [2000, 381–405]). Following the usual methodology of probabilistic modeling, the probabilities for various different kinds of elaborations could be based (at least as a first approximation) on counts of them in an analyzed corpus—for example, a set of published Schenkerian analyses. A study by Larson (1997–98) offers an interesting first step in this direction. Gilbert and Conklin (2007) also attempt to model hierarchical pitch structure with a probabilistic CFG.
39 Chomsky (1957, 22).

simpler ways, without requiring all the apparatus of a CFG. At local levels (for example, a tonally closed phrase or period), it could be modeled quite well with a finite-state model of harmonic progressions; indeed, it appears to be captured well by the model in Example 1. At larger levels, too, it is not clear what the *Ursatz* adds, in predictive power, to a more conventional model representing a piece as a series of key areas with a progression of chords in each key, beginning and ending in the same key, and ending with a perfect cadence in the home key.

In short, it is unclear what is to be gained by modeling Schenkerian theory as a CFG. To justify this approach, one would need to find "long-distance dependencies" that are not predicted by more conventional principles of tonal theory. I am not convinced that such long-distance dependencies exist. Rather than pursuing this approach any further, I will argue that there are other ways of testing Schenkerian theory (as a theory of compositional practice) that seem more likely to lead to successful results.

2.4 *ZÜGE* AND *STUFEN*

The linear progression or *Zug*—a stepwise line connecting two structural tones—is a ubiquitous element in Schenkerian analysis and is given great importance in Schenker's own writings. One might wonder if this concept could be used to make testable predictions about tonal music. The simple fact that stepwise melodic intervals are very prevalent in tonal melodies could be predicted by a simple note-to-note model, and does not require *Züge*. However, the *Zug* concept suggests a further prediction: If most stepwise motions are part of a larger *Zug*, this suggests that a step (whole step or half step) is likely to be followed by another step in the same direction. The idea that a small melodic interval tends to be followed by another same-direction interval—sometimes called "process" or "inertia"[40]—has been widely discussed in connection with melodic expectation (I will return to this below). As a prediction about compositional practice, the inertia idea can be tested quite easily, using computationally encoded corpora. Example 5 shows the results for two corpora, one of Bach chorales and one of Haydn and Mozart string quartets. In both cases, the prediction is strongly confirmed: an ascending step is much more likely to be followed by another ascending step than by a descending step, whereas for descending steps, the reverse is true.[41]

An extension of the *Zug* idea—and one that is perhaps more distinctively Schenkerian—is that linear progressions operate not just at the surface of melodies but at higher levels. In ST, a melody is typically analyzed as a long-range linear progression, recursively elaborated by other linear progressions (along with other elaborations such as arpeggiations). However, this does not appear to constitute a very strong prediction. It is true that

40 See Narmour (1990, 2-4) for "process" and Larson (1997, 102) for "inertia."

41 Huron (2006, 77–78) reports a similar study using a large corpus of both Western and non-Western melodies; the data reflect "step inertia" for descending steps only, not ascending steps. This intriguing result suggests that the bi-directional step inertia found in common-practice music may be peculiar to that style, or at least not universal.

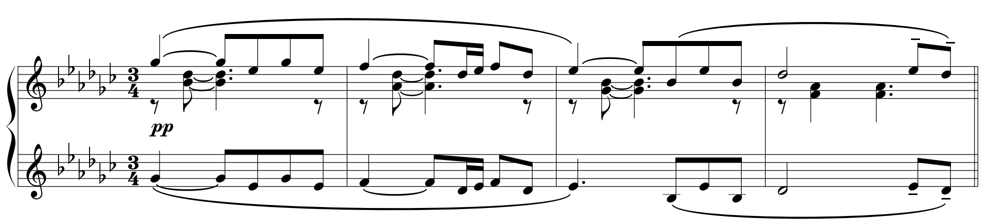
Bach chorales First interval	Second interval Ascending	Descending
Ascending	70.0%	30.0%
Descending	32.3%	67.7%
Mozart / Haydn string quartets First interval	Second interval Ascending	Descending
Ascending	64.4%	35.6%
Descending	30.7%	69.3%

EXAMPLE 5. *Stepwise intervals in Bach chorales and in Mozart and Haydn string quartets*
The table shows, for all cases of two successive steps within a melodic line, the proportion of cases in which an initial ascending or descending step (half-step or whole-step) was followed by an ascending or descending step. Two corpora were used: ninety-eight Bach chorales (in their entirety) and the first eight measures of all of Mozart's and Haydn's string-quartet movements. In both corpora, all four melodic lines were analyzed.

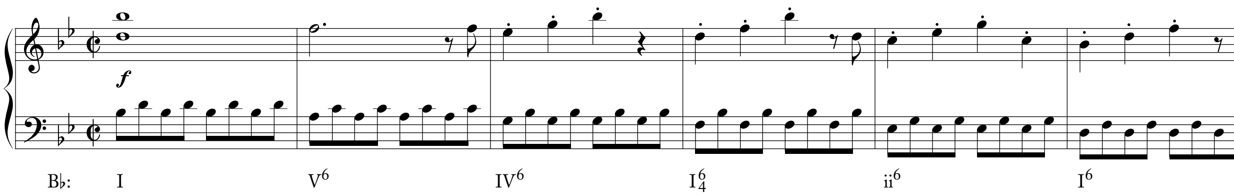
one usually can find a long-range stepwise linear progression across a melody; but this may largely be due to the fact that melodies are usually confined to a fairly limited pitch range. In many cases, the predictive power of linear progressions seems to lie not so much in pitch patterns themselves, but rather in the way that pitch patterns are realized rhythmically and metrically. As an especially obvious example, consider the melody "Un bel dì" from Puccini's *Madama Butterfly* (Example 6). Considered simply as a sequence of pitches, this melody could be generated in many ways; generating it hierarchically from the linear progression G♭–F–E♭–D♭ seems no more compelling than other alternatives. But when rhythmic context is added, we see that each of the notes G♭–F–E♭–D♭ falls on the downbeat and is agogically accented. This seems like more than mere coincidence and suggests to me that the composer was aware of the linear pattern in choosing a rhythmic setting for the melody; one might argue on this basis that he had the same linear framework in mind in generating the pitches as well.

Another way in which linear progressions might yield testable predictions concerns harmonic progressions. Consider the progression shown in Example 7, from the first movement of Mozart's Symphony No. 40 in G Minor, K. 550—a progression discussed by John Rothgeb.[42] As Rothgeb notes, viewed in conventional harmonic terms, this progression seems extremely odd: it contains root motions from V to IV and from ii to I, both of which are usually considered incorrect. And yet the progression does not sound wrong and is not uncommon in tonal music. (Two other examples are the opening of Handel's Concerto Grosso, Op. 6, No. 4, and the second theme of Brahms's Violin Sonata in G Major, Op. 78, I, mm. 40–42.) The Schenkerian perspective provides an explanation: the logic of the progression is built on a stepwise descending bass line. By

42 Rothgeb (1975, 277–79).



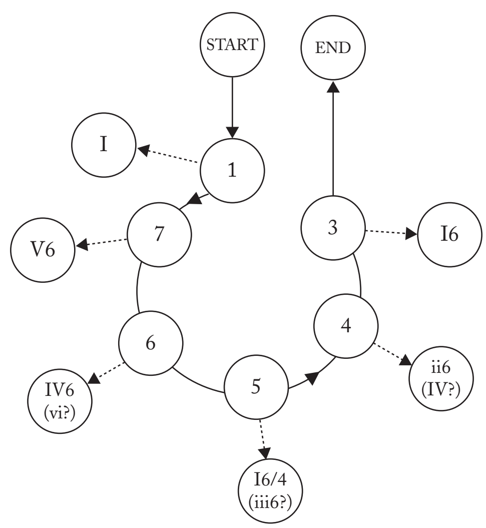
example 6. *Puccini, "Un bel dì" from* Madama Butterfly



example 7. *Mozart, Symphony No. 40 in G Minor, K. 550, I, mm. 28–33*

this view, the chords are not to be understood as constituting a conventional harmonic progression, but rather as a kind of elaboration of the underlying bass notes. Thus a linear progression can license harmonic "successions" (not really progressions at all) which would not normally be allowed.

A possible generative model of this process could be constructed in the following way: We could begin with a harmonic finite-state model such as that shown in Example 1, which generates a series of harmonies and then generates pitches from these harmonies. This could be augmented by an additional network that generates bass-line linear progressions and then forms triads on the bass notes. Upon entering a functional harmony node such as T, the machine would have the option of jumping into a linear-progression network and generating a *Zug* based on that harmony.[43] A fragment of such a model—showing just the part needed for the *Zug* in Example 7—is shown in Example 8. Many questions then arise. What are the constraints on when linear progressions can occur? By general agreement, linear progressions usually connect two notes of an underlying harmony; a strict view might hold that they can only arise this way. But one often sees linear patterns in tonal music that cannot easily be explained as prolongations of an underlying harmony—for example, a linear bass progression $\hat{1}$–$\hat{7}$–$\hat{6}$–$\hat{5}$–$\hat{4}$ in the first part of a phrase under what would normally be considered a prolongation of I (certainly not a prolongation of IV).[44] There are also constraints on the triads formed from the bass notes, which appear to be quite complex. In Example 7, for example, why does Mozart choose a I^6_4 triad over the F rather than (for example) a iii^6 triad?

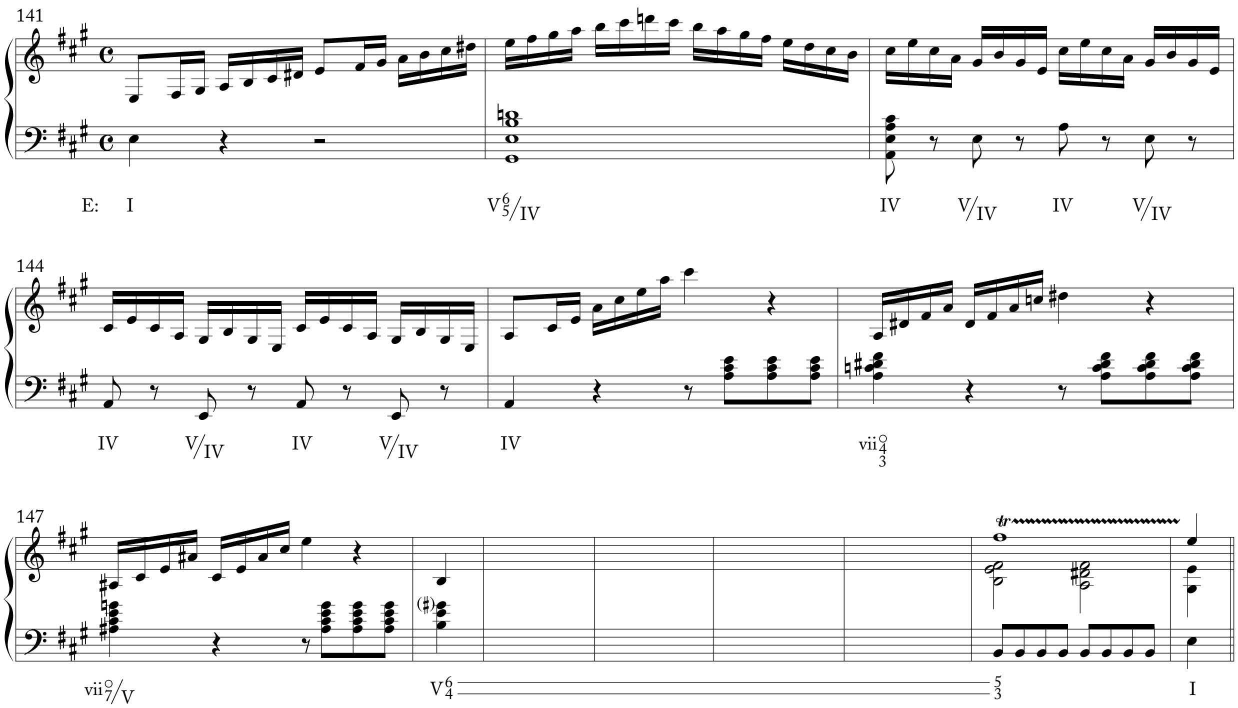


example 8. *A finite-state model for generating a* Zug

A further situation in which Schenkerian principles may make testable predictions is illustrated by Example 9, a passage from Mozart's Clarinet Concerto in A Major, K. 622, I. The progression of local chords, shown below the score, is mostly quite conventional, and could perhaps be accommodated by a finite-state model of harmony such as Example 1 (suitably expanded to include chromatic harmonies and applied chords). But in some respects it seems rather arbitrary: Why the alternations between IV and V/IV in mm. 143–44; why the diminished seventh chords in mm. 146–47? It can be seen that there is a large-scale logic governing this passage: the first part of the progression centers around IV and the second part around V, leading finally to a I in m. 154, thus forming a large-scale IV–V–I cadence. Rather than generating the progression as a series of local harmonic moves, it might be generated more convincingly by a

43 This is essentially an "Augmented Transition Network"—a formalism once fashionable in psycholinguistics (see Garman [1990, 345–55]).

44 Cadwallader and Gagné (2007, 74–75) suggest that linear progressions sometimes serve to connect one harmony to another, rather than prolonging a single harmony; they explain $\hat{1}$–$\hat{7}$–$\hat{6}$–$\hat{5}$–$\hat{4}$ progressions in this way.



EXAMPLE 9. *Mozart, Clarinet Concerto in A Major, K. 622, I, mm. 141–54*

more hierarchical process, which first constructed an underlying progression of *Stufen*, IV–V–I (in itself a very normative progression), and then generated surface chords as elaborations of these structural harmonies. Once mm. 142–45 are identified as an expanded IV, the surface progression of those measures (interpreted as I–V–I–V . . . in relation to IV) becomes extremely normative; the diminished-seventh chords in mm. 146–47 could be treated as linear connectives between the IV and the V (using the logic of linearly generated harmony discussed earlier). In this way, a hierarchical view of the passage's harmony makes it seem more normative, and thus predicts it more strongly than a purely finite-state model.

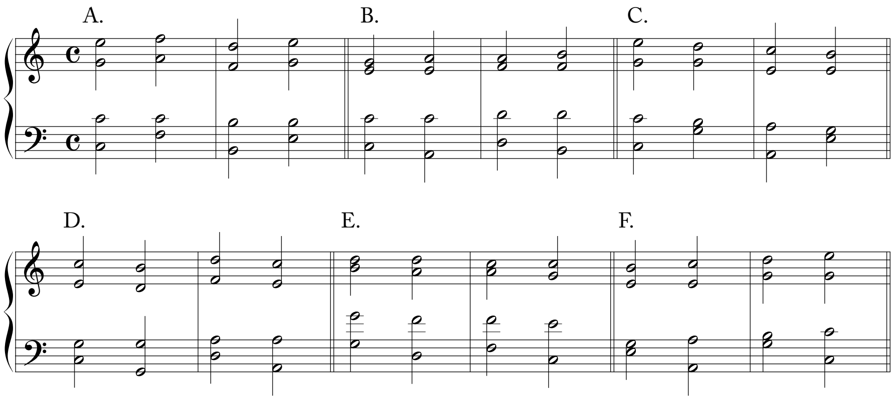
2.5 PREDICTING OTHER PHENOMENA OF TONAL MUSIC

So far my approach to evaluating ST as a theory of tonal composition has been to examine the theory and consider what its predictions might be. Another way of approaching this problem is to examine phenomena of tonal music—especially phenomena that have not been satisfactorily explained by other means—and consider how ST might predict them. Here I will consider just two aspects of the common-practice musical language for which, as far as I know, no adequate explanation has been provided.

As is well known, certain key schemes are especially common in tonal pieces. For one thing, given a particular main key (e.g., C major) certain secondary keys are especially common (such as G major, F major, and A minor). Various theories of key relations—often employing spatial representations—have been proposed that explain these facts fairly well.[45] What has not been explained so well is the temporal arrangement of keys, which also reflects consistent patterns. The key scheme I–V–vi–I is extremely common; it is frequently seen, for example, in Baroque suite movements and classical-period minuets and slow movements. Other patterns such as I–V–IV(ii)–I or I–V–vi–IV(ii)–I are also common. What is especially puzzling about these patterns is that they do not follow conventional rules of harmonic progression; indeed, they are virtually mirror images of conventional progressions. I–V–IV–I is a common key scheme but an uncommon harmonic progression; I–IV–V–I is ubiquitous as a harmonic progression, but very rare as a key scheme.

Could ST be of any value to us in explaining these patterns? Here we encounter the difficult issue of the relationship of Schenkerian structure to what might be called "key structure"—the arrangement of key sections within a piece. Schenkerian analysis does of course feature tonal entities prolonged over large portions of a piece, and a naïve encounter with a Schenkerian analysis might suggest that these entities corresponded to key sections. It should be clear, however, that this is not the case. Sometimes, high-level events in a Schenkerian analysis do represent the tonic triads of key sections; but very often they do not.[46] If one considers the deepest structural level of a Schenkerian reduction—the *Ursatz*—the V triad almost never

45 Krumhansl (1990, 40–49); Lerdahl (2001, 41–88).
46 For an insightful discussion of this point, see Schachter (1987, 292–96).



EXAMPLE 10. *Common and uncommon sequential patterns*

represents an extended section in the dominant key, and indeed is often not tonicized at all. Conversely, a key may be implied without its tonic triad being represented as a high-level reductional event; and when the tonic triad is a high-level event, the span of music over which it is prolonged may be quite different from that of the corresponding key section.

In short, Schenkerian analysis does not appear to represent key sections in any direct way. Even if key sections did consistently correspond to *Stufen*, the prediction that would seem to follow from this—if any—is that key schemes follow the same logic as harmonic progressions; and as we have seen, this is far from the case. It is true that key areas do sometimes correspond to *Stufen*, and that treating key sections as expanded harmonies sometimes helps to make sense of them. Example 9 offers a case in point: the tonicization of IV is nicely explained by treating it as an expanded predominant of a large structural cadence. But for the most part, the correspondence between *Stufen* and key sections does not hold. It remains possible, however, that ST could be used to generate predictions about key structure in less direct way—for example, by specifying which *Stufen* permit (or require) tonicization and which ones do not. This possibility has not been much explored, but it might offer an interesting way of testing ST as a model of compositional practice.

Another aspect of the common-practice language that has not been fully explained is sequences—melodic patterns that repeat at changing pitch levels. While sequences could conceivably occur in all kinds of different harmonic contexts, it is well known that certain harmonic patterns are particularly common and account for nearly all sequences in the common-practice repertory. The patterns most often presented in theory textbooks are the first four shown in Example 10. Patterns A, B, and C are all widely used; pattern D, the ascending-fifth sequence, sometimes occurs but is much less common than the descending-fifth pattern (A).[47] In some respects, sequences are well-suited to Schenkerian analysis. Since the melodic patterns of sequences generally shift up or down by a step at each repetition (with the exception of pattern C, which shifts by a third), they can often be analyzed as elaborations of linear progressions (though it is sometimes uncertain which notes of the pattern are "structural"). Some authors have argued that sequences are best understood as being essentially contrapuntal, rather than harmonic, in origin. Allen Cadwallader and David Gagné, for example, refer to sequences as "linear intervallic patterns" and write:

> In textures of more than two voices, chords naturally arise in conjunction with the repeated pattern, thereby forming chordal sequences. The chords in the pattern . . . therefore result from contrapuntal motion.[48]

This view is also reflected in contrapuntal names for sequences such as "ascending 5–6" for pattern B and "descending 5–6" for pattern C.[49]

From the point of view of explaining compositional practice, the essential question posed by sequences is this: Why is it that certain patterns are widely used and others are not? One could try to explain these regularities in contrapuntal terms. The popularity of outer-voice patterns such as 10–6 or 10–10 reflects the preference for imperfect consonances in species counterpoint; the alternating 5–6 pattern is also familiar from species and can be seen as a strategy to avoid parallel fifths. However, the contrapuntal view of sequences only takes us so far. In particular, it fails to explain why patterns A, B, and C are all much preferred over their mirror images, patterns D, E, and F. These mirror-image patterns seem contrapuntally just as good as those of A, B, and C; why are they so rarely used? (As noted earlier, the ascending-fifth pattern is generally described as much less common than the descending-fifth pattern.) A more promising explanation for these facts may lie elsewhere. Traditional harmonic theory holds that descending fifth motions are preferred over ascending fifths, and descending thirds over ascending thirds;[50] this nicely predicts the preference

47 Aldwell and Schachter note that "most diatonic chordal sequences fall into one of [these four] categories" (2003, 265); see also Laitz (2008, Chapter 22). On the rarity of the ascending-fifth sequence, see Aldwell and Schachter (2003, 248); and Laitz (2008, 501).

48 Cadwallader and Gagné (2007, 81–82).

49 Aldwell and Schachter (2003, 273–78).

50 See, for example, Schoenberg ([1954] 1969, 6–8).

EXAMPLE 11. *(a) Beethoven, Symphony No. 6 in F Major, Op. 68, I, mm. 9–12; (b) Measures 117–20 from the same movement; (c) A possible reduction of (b); (d) Measures 13–16 from the same movement*

for pattern A (all descending fifths) over D (all ascending fifths), and the preference for B (descending fifths and thirds) over E (ascending thirds and fifths). The predictions of harmonic theory regarding C (ascending fifth/ascending step) and F (descending fifth/descending step) are less successful. But harmonic theory would seem to present at least a partial explanation for the preferred sequential patterns; the contrapuntal view of sequences does not. Here again, then, is a puzzling set of facts that a "theory of tonal music" might reasonably be expected to explain, yet ST seems to be of little help to us.

Overall, the view I have presented of ST as a theory of composition is fairly pessimistic. Formulating the theory as a CFG seems to offer little hope of making substantive predictions about tonal music. I also considered certain phenomena about tonal music that one might like a theory to predict, such as regularities of key structure and common sequential patterns; ST seems to have little to say about these phenomena (though perhaps future extensions of the theory will shed light on these issues). However, some aspects of Schenkerian structure *do* yield clear and plausible predictions, notably certain fairly local uses of *Züge* and *Stufen*; and these aspects of the theory might make an important contribution to the modeling of tonal music. I will return at the end of the article to what all this means for our view of ST. I now turn to the second issue posed at the beginning of this essay: the testing of ST as a theory of perception.

3. SCHENKERIAN THEORY AS A THEORY OF PERCEPTION

The hypothesis under consideration here is that Schenkerian structures are formed in the minds of listeners as they listen to tonal music. One basic question for this construal of the theory is: Which listeners does the theory purport to describe? A theory of listening might claim to apply, for example, to all those having some familiarity with tonal music, or perhaps only to those with a high degree of musical training. Given the speculative nature of this discussion, there is no point in trying to be very precise about this. My general focus here—and the focus of most of the relevant experimental research as well—is on listeners who have some musical training and familiarity with common-practice music, but who have not had extensive training in ST.[51] This is not in any way to deny the validity of perception informed by advanced training in music theory; but to claim something as a "theory of perception" implies that it characterizes people's perception beyond what is due to explicit study of the theory.[52]

A natural approach to testing ST as a theory of perception is to use the methods of experimental music psychology. Rather than simply asking listeners to report their mental representations of a piece, the usual approach of music perception research is to access these representations in an indirect fashion, usually by having subjects make judgments about some aspect of musical passages such as their similarity, tension, or expectedness. In what follows I will review some of this research; I will also add some of my own thoughts as to the aspects of ST that seem most promising with regard to the modeling of perception.

3.1 SIMILARITY

One way of testing the perceptual validity of ST is through judgments of similarity. If people form reductions of pieces as they listen, it stands to reason that passages sharing the same underlying structure should seem more similar than those that do not. This reasoning has sometimes been used to argue for the perceptual reality of pitch reduction. For example, Fred Lerdahl and Ray Jackendoff point to the two melodic passages from Beethoven's *Pastoral* Symphony shown in Example 11, and note that listeners have no trouble hearing Example 11(b) as an elaboration of Example 11(a); they argue that such judgments are evidence that the second melody is understood as an elaborated

51 Most of the experimental studies described below, such as Serafine et al. (1989) and Bigand and Parncutt (1999), used undergraduate students with varying amounts of musical training.

52 For discussion of this point, see Temperley (1999, 75–76).

version of the first (or, perhaps, that both are heard as elaborations of the same underlying structure).[53]

The possibility of testing the psychological reality of pitch reduction through similarity judgments was explored in a study by Mary Louise Serafine et al.[54] In one experiment, the authors took passages from Bach's unaccompanied violin and cello suites, generated correct and incorrect reductions for them, and had subjects judge which of the two reductions more closely resembled the original. Subjects judged the correct reductions as more similar at levels slightly better than chance. As Serafine and her collaborators discuss, one problem with this methodology is that there are many criteria by which passages might be judged as similar, some of which are confounded with Schenkerian structure. Given two reductions of a passage, if one reduction is inherently more musically coherent than the other, it may be judged more similar to the original passage simply because it is more coherent; this in itself does not prove the perceptual reality of reduction. Another possible confound is harmony. Two passages based on the same surface harmonic progression are likely to seem similar, and a passage and its Schenkerian reduction are also likely to share the same implied harmonic progression; thus any perceived similarity between them may simply be due to their shared harmony.

Serafine et al. made great efforts to control for these confounds. With regard to the "coherence" problem, they had subjects indicate their degree of preference for "correct" and "incorrect" reductions heard in isolation; these judgments were not found to correlate strongly with the perceived similarity between the reductions and surface patterns. The authors also tried to tease out the effects of surface harmony from Schenkerian structure. In one experiment, listeners compared passages with "harmony foils" that were different from the original passage in surface harmony but alike in Schenkerian structure, and "counterpoint foils" that were different in Schenkerian structure but alike in harmony. After a single hearing, subjects tended to judge the harmony foils as more similar to the originals than the counterpoint foils; after repeated listening, however, they tended to judge the counterpoint foils as more similar, suggesting (somewhat surprisingly, as the authors note) that Schenkerian structure is a more important factor in similarity on an initial hearing but that surface harmony becomes more important with repeated hearings.[55]

The work of Serafine et al. is a noble attempt to explore the perceptual reality of ST using experimental methods. The methodological pitfalls of using similarity judgments have already been discussed; I wish to add a further note of caution about this approach. Let us return to the *Pastoral* excerpts in Example 11. If someone unfamiliar with the piece were simply given Example 11(b) and asked to indicate (by singing, for example) what underlying melody it was an elaboration of, I wonder if they would produce Example 11(a). They might well judge Example 11(a) as a plausible reduction of Example 11(b), if this were suggested; but other melodies might also be considered plausible reductions, such as Example 11(c). One might argue that the ease of hearing the connection between Examples 11(a) and 11(b) in context demonstrates the remarkable ability of listeners to find connections between two melodies that are presented side by side (or presented in formally parallel positions—that is, in positions in which one expects a theme to be repeated). No doubt, this process often involves a kind of reduction. The question is whether listeners perform such reductions spontaneously even when such a juxtaposition is not involved. Comparing one musical pattern to another, and extracting their similarities and differences, may often involve a simplification or reduction of the more complex pattern. But this does not mean we spontaneously reduce complex patterns to simpler ones. And the fact that many different simplifications of a complex pattern—not all of which, presumably, are generated spontaneously—often seem like plausible reductions of it suggests that such intuitions are not particularly good evidence of spontaneous reduction.

Notwithstanding all this, it seems clear that in many cases, listeners do perform reduction—as evidenced by their ability to recognize the similarity between a melody and an elaborated version of it. Such situations are of course commonplace in common-practice music; they include not only actual theme-and-variations pieces, but the countless other situations in which a theme is followed by an elaborated repeat, such as the *Pastoral* melody in Examples 11(a) and (b). (Romantic piano music—especially Chopin's—also offers innumerable examples of this.) The reductive process involved here seems to be highly susceptible to the "power of suggestion"; that is, our reduction of a passage in listening may be steered toward whatever has been presented as the simplified version of it. And yet, not *everything* is a plausible reduction; for example, the second half of the Beethoven theme, shown in Example 11(d), is not a plausible reduction of the variation melody in Example 11(b). An interesting question is then, what are the factors that make one pattern a plausible reduction of another—or, to put it differently, what are the criteria whereby reduction is done? A number of factors come to mind here, all well known from tutorials on Schenkerian reduction[56] and related work.[57] In general, a plausible reduction of a passage retains the notes that are rhythmically emphasized (through duration and metrical placement) and harmonically stable (chord-tones rather than non-chord-tones), and also makes some kind of inherent musical sense, with the notes of the reduction forming coherent, mostly stepwise lines.

53 Lerdahl and Jackendoff (1983, 105).

54 Serafine et al. (1989).

55 A similar study by Dibben (1994) tested the perceived similarity between note patterns and correct or incorrect reductions, and found that correct reductions were perceived as more similar than incorrect ones. This study used Lerdahl and Jackendoff's (1983) time-span reduction rather than Schenkerian reduction, so it is less relevant for present purposes; time-span reductions are rather different from Schenkerian ones, as Dibben acknowledges.

56 Beach (1983); Cadwallader and Gagné (2007).

57 E.g., Lerdahl and Jackendoff (1983).

To explore these criteria experimentally would seem to be an interesting project for music psychology—a project already begun by the work of Serafine et al. cited above.

One might also use similarity judgments as evidence of reduction in a more indirect way. If two passages are reduced in similar ways, then one might expect them to have similar effects or functions with respect to other aspects of music cognition. Example 9, the Mozart concerto passage discussed earlier, is a case in point. I argued that this passage can be seen to express a large IV–V–I cadence; it seems reasonable to suppose that an experienced listener of tonal music would comprehend this, and thus would experience a sense of large-scale closure at this point. But this large-scale progression cannot be understood unless one reduces the surface harmonies. If listeners did indeed comprehend the closural effect of this passage, this could be taken as evidence that they perceived the underlying cadence and thus were performing reduction; this might be an interesting area for experiment.

3.2 TYPICALITY AND EXPECTATION

Listeners can, to some extent at least, judge the typicality or normality of a piece of music, in relation to the norms of a style with which they are familiar. A classically trained listener can judge that a string quartet by Mozart is more normative within the classical style than one by Elliott Carter. Such judgments form a kind of psychological data that can be used to test theories of music perception. With regard to ST, the obvious prediction is that passages judged as normative by the theory will also be judged as such by listeners. This raises the question of what kinds of patterns the theory judges as normative. One obvious candidate is the *Ursatz* itself, which is claimed to be the background structure not only for complete pieces but also for smaller tonally closed formal units such as periods and sections. This suggests the prediction that a short tonal piece constructed on an *Ursatz* should be judged as more normative than one that is not.[58] As with similarity judgments, the challenge in constructing such passages would be to avoid other confounding factors—in particular, tonal closure itself. If a conventional *Ursatz*-based passage were compared to one that was not tonally closed at all (i.e., not ending on V–I in tonic), no doubt listeners would judge the first as more normative, but that could be easily explained in more conventional harmonic terms (e.g., using finite-state harmonic models such as Example 1). One could focus instead on the melodic aspect of the *Ursatz*, comparing two passages based on the same normative harmonic structure but with different melodies, one that featured a clear *Urlinie* (such as $\hat{3}$–$\hat{2}$–$\hat{1}$) and one that did not. Of course, many actual pieces and tonally closed sections do not feature clear *Urlinie*; such passages are typically analyzed in ST either by locating the *Urlinie* in an inner voice or by indicating one or more of its notes as "implied."[59] But one might argue that the theory judges such structures as less normative than those with explicit, top-voice *Urlinien*, and that listeners should therefore do so as well.

A related aspect of perception deserving consideration here is expectation. A great deal of work in music perception has focused on modeling listeners' expectations.[60] In most such studies, subjects are played some kind of context followed by various possible continuations, and must then indicate the degree to which the continuation seems "good" or expected. Again, the prediction of ST is that the continuations expected by listeners will be those judged as normative by the theory. One important body of expectation data is shown in Example 12, taken from a study by Lola Cuddy and Carol Lunney.[61] In this study subjects were played a single "context interval" (two notes in succession) followed by a "continuation tone," and had to judge the expectedness of the continuation tone given the context interval. The figure shows Cuddy and Lunney's results for context intervals of descending major sixth and ascending major second. To a large extent, these results can be explained by basic factors of scale membership (pitches within the key or keys implied by the context interval are more expected) and pitch proximity (pitches close to the second pitch of the context interval are more expected); the principle of "reversal"[62] is also apparent, as the most expected pitches after the major sixth involve a change of direction.[63]

One interesting feature of Cuddy and Lunney's data concerns the phenomenon of "inertia." As noted earlier, the centrality of linear progressions in ST suggests that melodic steps should tend to be followed by further steps in the same direction; if the theory applies to perception, then melodic expectancies should reflect this tendency as well. Inspecting Cuddy and Lunney's data for the ascending-second context, we see that continuation intervals of +1 and +2—an ascending half-step and whole-step—are indeed highly expected, more so than descending steps. Notably, this is not predicted by any of the principles discussed in the previous paragraph: key membership, pitch proximity, or reversal. The fact that +3 (an ascending minor third) is also highly expected—even more than +1 and +2—does not accord so well with ST. But there seems to be at least some evidence here of "inertia" in melodic expectation judgments, suggesting an awareness of linear progressions.

Schenkerian theory then suggests a further question: Do listeners also project expectations of linear progressions at larger structural levels? Consider the progression in Example 7. It was noted that harmonic motions such as V^6–IV^6 are generally rare

58 It seems reasonable to assume that the psychological validity of Schenkerian structures, if any, declines over longer time spans. Indeed, this has been found to be true of tonal closure itself; Cook (1987b) found that, with all but the shortest pieces, listeners (even those with considerable musical training) were unable to tell whether the piece ended in the same key in which it began.

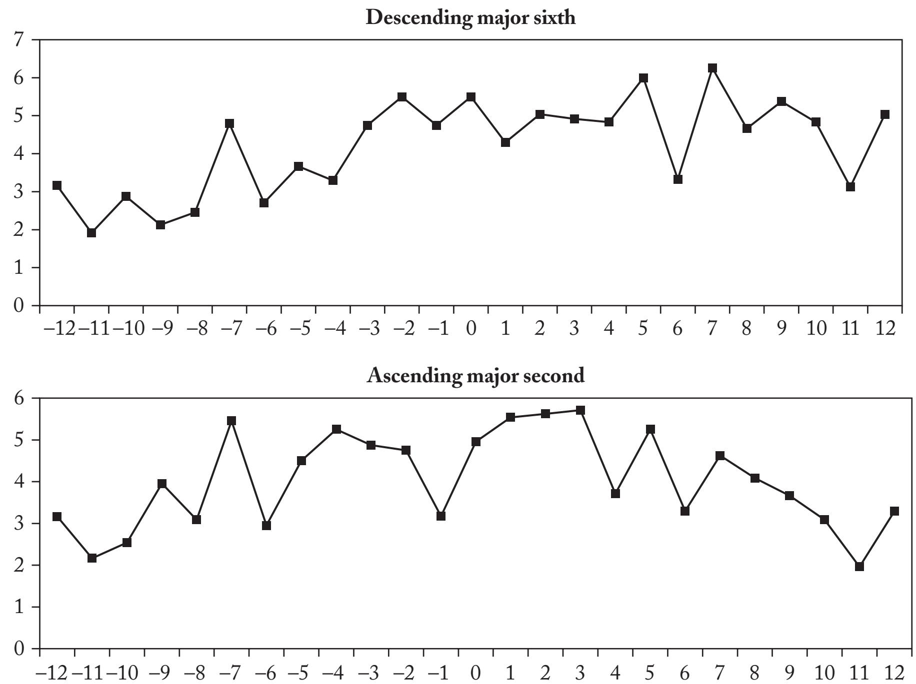
59 See, for example, Schenker ([1935] 1979, 51); and Cadwallader and Gagné (2007, 21–23).

60 See, for example, Schmuckler (1989); Narmour (1990); Cuddy and Lunney (1995); and Schellenberg (1996).

61 Cuddy and Lunney (1995, 460–61).

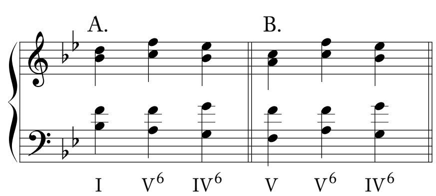
62 Narmour (1990, 5–6).

63 However, see Von Hippel and Huron (2000) for an alternative explanation of this phenomenon.



example 12. *Melodic expectation data from Cuddy and Lunney (1995)*
The data show average expectedness ratings (on a scale of 1 to 7) for continuation tones given context intervals of a descending major sixth (A_4–C_4) and an ascending major second ($B\flat_3$–C_4). The continuation tones are shown in relation to the second tone of the context interval; for example, –1 represents a continuation tone of B_3.

in tonal music, but in a context such as Example 7, they seem to be justified by the linear progression in the bass. This could be tested perceptually: does a IV^6 following a V^6 seem more expected or appropriate in a linear-progression context, such as Example 13(a), as opposed to another usage of V^6 such as Example 13(b), where no linear progression is involved? To my ears at least, there is no question that V^6–IV^6 seems "righter" in a linear progression context. This is a kind of testable prediction that might allow us to evaluate ST, or at least one aspect of it, as a theory of music perception.



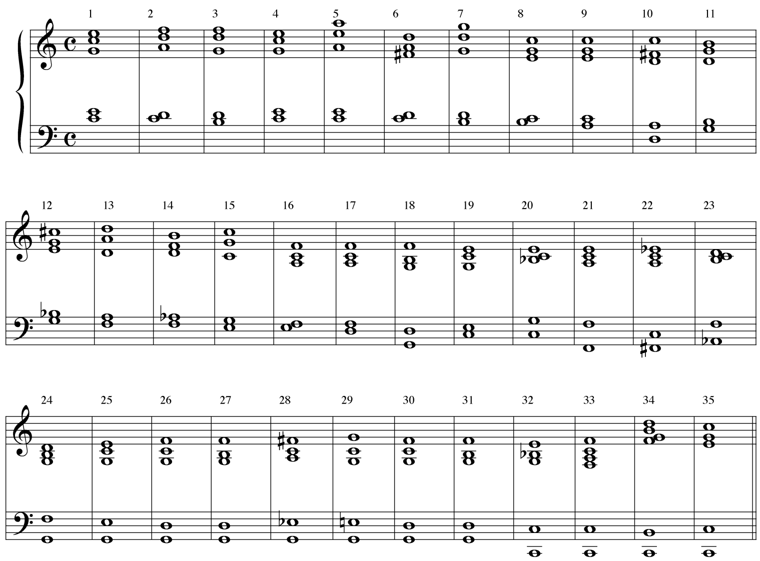
example 13. *V^6–IV^6 progressions in two different contexts*

3.3 motion and tension

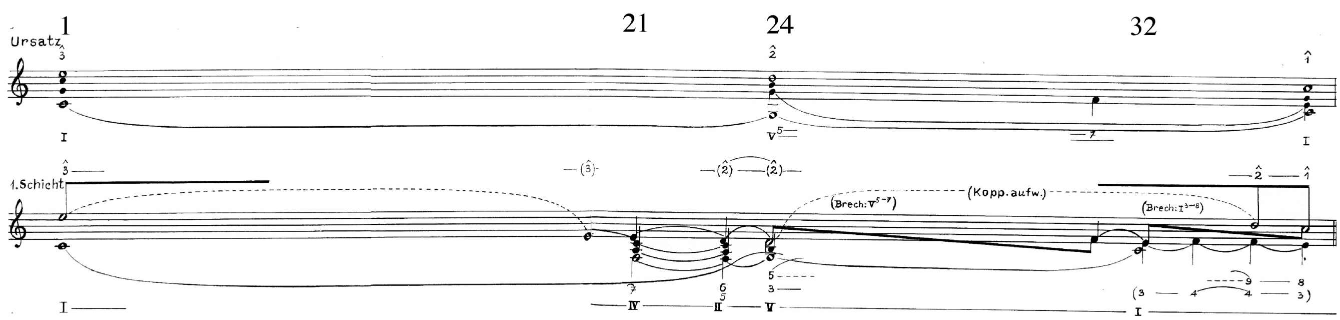
Another way of exploring the role of ST in perception is through its effect on the experience of musical motion. A number of authors have extolled the theory's power to explain this aspect of musical experience. Almost sixty years ago, Milton Babbitt heralded ST as "a body of analytical procedures which reflect the perception of a musical work as a *dynamic* totality."[64] Felix Salzer, in critiquing a Roman-numeral analysis of a Bach prelude, laments, "What has this analysis revealed of the phrase's motion, and of the function of the chords and sequences within that motion?";[65] the clear implication is that ST fulfills this need. Similarly, in Nicholas Cook's introduction to ST in *A Guide to Musical Analysis*, he argues that the theory—unlike Roman numeral analysis—shows how each chord in a piece is "experienced as a part of a larger motion *towards* some future harmonic goal."[66] But how does a Schenkerian analysis map onto experiences of motion; what exactly are the theory's predictions in this regard? Very few concrete discussions of this issue can be found. One of them is Cook's analysis (in the essay cited above) of Bach's Prelude in C Major from *The Well-Tempered Clavier*, Book I; thus we will focus on it at some length.

Cook's discussion of the prelude is framed very much as an investigation of where its essential motion resides. (Example 14 shows a slight reduction of the piece, with rhythmic figuration removed.) Cook notes first that the "beginning of the end" of the piece is the move to V at m. 24, and that the "end of the beginning" is m. 19—the end of a large-scale prolongation of I. (The tonicization of V in mm. 5–11 is not "structural"; thus the

64 Babbitt (1952, 262); italics mine.
65 Salzer (1952, 1: 11).
66 Cook (1987a, 29), italics in original.



EXAMPLE 14. *Bach, Prelude in C Major, from* The Well-Tempered Clavier, *Book I*



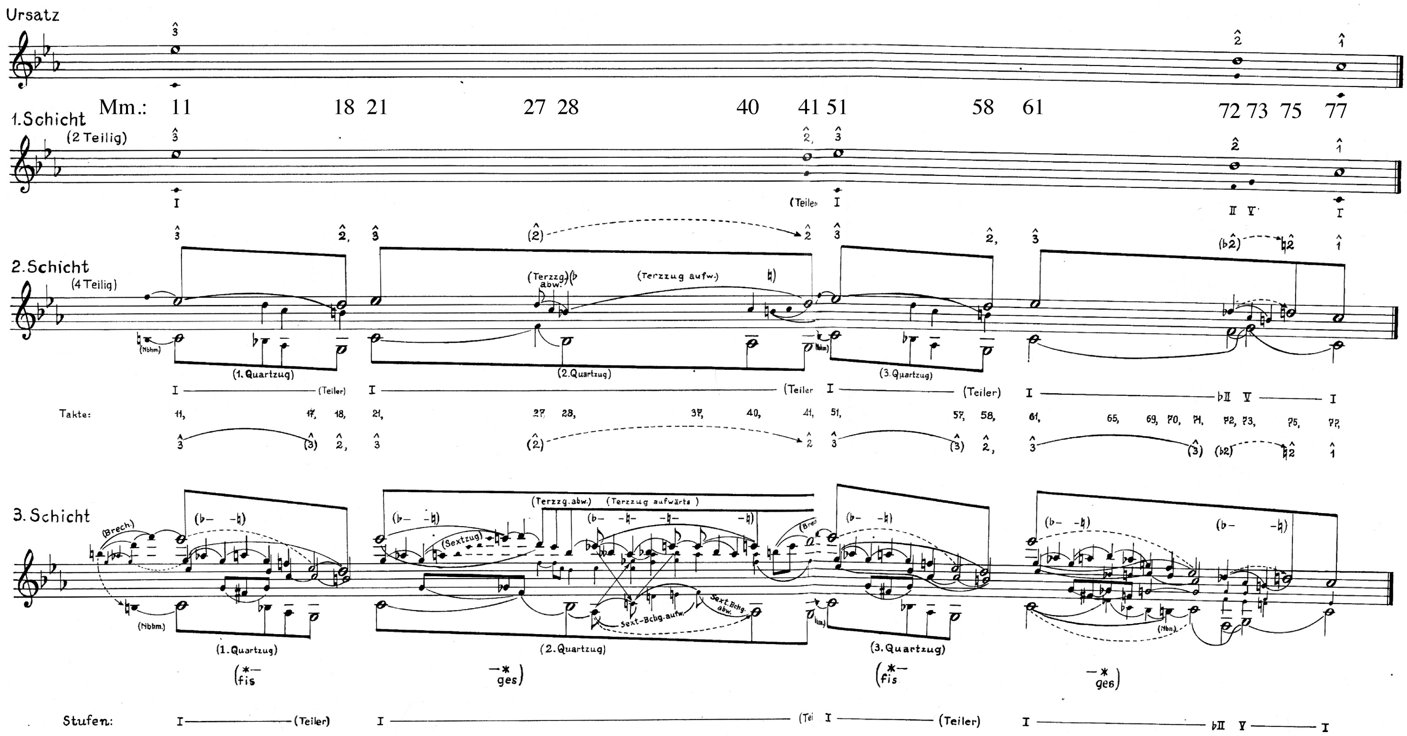
EXAMPLE 15. *Schenker's middleground and background analysis ([1933] 1969) of Bach's Prelude in C Major, from* The Well-Tempered Clavier, Book I

I at m. 19 is "continuous" with the initial I and not a "return.") It is not until m. 21 that we experience a "structural departure" from the initial tonic. Thus mm. 20–23 constitute the essential motion of the piece: "Harmonically, the entire evolution of the piece . . . lies in these four bars."[67] Cook then goes on to show how Schenker's analysis of the piece (shown in Example 15) brings out the contrast between the "tonally enclosed" beginning and ending sections and the goal-directed passage in between. The general point, then, seems to be that the *Ursatz* implies a motion from its initial I to its final I.[68] This does not, of course, imply that the "tonally enclosed" portions of the piece are entirely devoid of motion; indeed, Cook also discusses linear motions within the initial prolonged I. Presumably the pattern of stasis-motion-stasis within the *Ursatz* can be replicated at lower levels—wherever a *Stufe* is expanded with points of stability at either end and instability in between. But the implication is that the main motion of a tonal piece—the primary "harmonic evolution"—is that of the *Ursatz* itself, in the region just preceding the final cadential I, and that other motions are subsidiary.

Cook's account presents a hierarchical structure of motions of varying levels of importance, corresponding to levels of a Schenkerian reduction. This is a perfectly coherent proposal, and in the case of the Bach prelude it is not unreasonable. But as a general prediction about motion in tonal music, it seems rather counterintuitive. In most Schenkerian analyses (including Schenker's own), the V of the *Ursatz* and the immediately preceding predominant chords—corresponding to mm. 20–23 of the Bach prelude—occur at the very end of the piece (or perhaps just before the coda), so that almost the entire work is a prolongation of the initial I. For example, consider Schenker's analysis of Chopin's Etude Op. 10, No. 12, shown in Example 16.

67 Ibid. (32–34).

68 The logic here is actually not completely clear. According to Schenker's analysis, the IV–II harmonies of mm. 20–23 are not literally part of the *Ursatz*; but they are (perhaps) transitional between the initial I and the V of the *Ursatz* and therefore might be considered part of the "top-level" motion of the piece. Another question is whether the V of the *Ursatz* is part of this top-level motion as well; Cook seems to assume not, but it is not clear why.



EXAMPLE 16. *Schenker's middleground and background analysis ([1933] 1969) of Chopin, Etude in C Minor, Op. 10, No. 12*

According to the view that Cook seems to advocate, the first seventy-one of the eighty-four measures of this piece are characterized by relative stasis; it is only in m. 72 (or perhaps mm. 72–76, if the V of the *Ursatz* is included) that the main motion of the piece occurs. I would argue, rather, that the primary tonal motion of the piece is the motion away from the tonic key at m. 24, through the keys of B♭ major, F minor, and other fleeting tonalities, and back to the tonic at the "recapitulation" in m. 49.[69] The function of the II–V–I progression in mm. 72–77 is not to create large-scale tonal motion but rather to confirm the tonal destination at which we arrived some measures earlier. Intuitions about musical motion are obviously very subjective, and we should be cautious about projecting our own experiences on to other listeners. I can only say that, by *my* intuitions, the view of tonal motion advocated by Cook does not ring true.

The pattern of musical motion I have asserted for the Chopin etude might be predicted much more straightforwardly by a conventional key-based view of motion, in which the essential motion of a piece begins when the tonic key is left and ends when it is regained. I would not claim, however, that such a view is completely satisfactory either. One serious problem with the key-based account concerns large-scale tonal returns in sonata movements and other large forms. I would argue that the big "moment of return" in a sonata-form movement generally coincides with the return of the opening theme at the beginning of the recapitulation; importantly, this is not only the moment of *thematic* return (that much seems obvious) but of *tonal* return as well. But the return to the tonic *key*—as represented by a conventional Roman-numeral analysis—does not usually occur immediately at the return of the main theme, but rather somewhat earlier; the recapitulation is almost always preceded by a prolonged dominant which is clearly V/tonic (or V^7/tonic), not I/dominant. (In the case of the Chopin etude, the return to C minor occurs at m. 41 at the latest—the harmony here is V/C, not I/G; it might even be said to occur a bit earlier.) One could argue that the moment of return coincides with the first tonic harmony in the tonic key, but this argument, too, encounters problems. In many cases, the return to the tonic key even precedes the retransitional dominant, and may involve fleeting tonic harmonies.[70] A Schenkerian view could be of help here; in such an analysis, the return to I at the beginning of the recapitulation is usually quite a high-level event. But it is not usually a top-level event—it might be at the second or third level, as it is in Example 15—and it remains to be explained why the essential tonal motion of a piece should be conveyed at this level rather than at the top level.

69 The piece also contains a shorter tonal excursion in mm. 65–70 of the coda—short but dramatic, as it arguably moves further from the tonic than anything in the work's main tonal "journey."

70 Beethoven's Piano Sonata in C Major, Op. 53, I, offers a typical example. The thematic return—and surely the tonal return as well—occurs at m. 156. But this is preceded by a passage of some twenty-six measures in C (though mostly minor rather than major), including a very prominent C-minor harmony at mm. 132–33.

Another area of musical experience in which ST may make testable predictions is patterns of tension and relaxation (or similarly—perhaps equivalently—stability and instability). Of particular interest here is the work of Lerdahl.[71] Lerdahl's approach to musical tension builds on Lerdahl and Jackendoff's theory of "prolongational reduction," in which the notes of a piece are assigned a hierarchical structure quite similar to Schenkerian structure.[72] Lerdahl's theory assigns a tension value to each event in a prolongational reduction. The theory relies heavily on a spatial representation theory of chords and keys (which ST of course does not). Oversimplifying somewhat, the tension value of an event in a reduction is the "pitch space distance" between the event and its immediately superordinate event in the tree, plus the distance between the parent event and its superordinate event, and so on, up to the highest-level event in the piece. As a result, events that are more embedded in the tree carry higher tension, as do events that are less closely related to their contexts (such as chromatic notes and chords).

Lerdahl's tension theory has been the focus of several experimental studies. In a study by Emmanuel Bigand and Richard Parncutt,[73] subjects were instructed to judge the tension of chords at each point in a harmonic progression, and the results were compared to the predictions of Lerdahl's theory. At global levels, the match between the theory's predictions and tension judgments was poor; a phrase in a non-tonic key is rated high in tension by Lerdahl's theory, but was not judged to be so by listeners.[74] At more local levels, the theory predicted tension judgments fairly well, although the authors observe that this may be largely due to the effect of cadences, which both the subjects and the theory judged to be low in tension. Experiments presented by Lerdahl and Carol L. Krumhansl,[75] in which the hierarchical tension model was applied along with melodic ("attraction") and surface dissonance factors, yielded more positive results; in a wide variety of musical excerpts, tension judgments were found to correlate strongly with the theory's predictions of hierarchical tension.

Altogether, the evidence gives at least partial support to Lerdahl's model as a predictor of perceived tension; there seems reason to hope that ST might yield successful results in this regard as well. An important question here concerns the role played by pitch space in Lerdahl's theory. Consider a chromatic chord in the middle of a phrase, which presumably would normally be perceived as high in tension. Lerdahl's theory assigns high tension to such a chord for two reasons: first, because it is deeply embedded in the reduction; secondly, because it is harmonically remote from neighboring events. The first of these factors can be captured by ST; the second cannot. A skeptic might argue that the perceived tension of a chromatic chord is primarily due to the second factor—the simple fact that it is chromatic in relation to the current key; this could be represented without any use of reductional structure. Whether perceived tension is due more to pitch-space or to reductional factors is currently an open question. In any case, the possibility of testing ST as a perceptual theory through its predictions of tension deserves further study.

4. CONCLUSIONS

In this essay, I have explored the prospects for testing ST both as a theory of composition and a theory of perception. As a theory of composition, I suggested that we could evaluate ST by considering how well it predicts what happens and does not happen in tonal music. I first considered the idea of formulating the theory as a context-free grammar, but expressed skepticism as to the prospects for this enterprise. I then considered some specific concepts of ST, *Stufen* and *Züge*, and suggested some ways in which they might indeed make successful predictions about tonal music.

With regard to perception I considered several ways that the theory could be construed to make testable predictions about the processing of tonal music. ST predicts that passages with similar reductions should seem similar; it predicts that events judged as normative by the theory will be highly expected; and it may also make predictions about the experience of musical motion and tension. It seems to me that the prospects for the theory here are extremely mixed. Listeners' ability to perceive the similarity between a theme and a variation of it seems to point to some kind of reduction; but we cannot conclude from this that listeners spontaneously generate reductions as they listen. The potential for the theory to tell us much about musical motion also seems doubtful, or at least unproven. On the other hand, melodic expectation data seem to indicate some awareness of linear progressions; reduction may also play a role in explaining tension judgments, as suggested by tests of Lerdahl's theory.[76] I also pointed to several specific situations—regarding linear harmonic progressions such as Example 7 and large-scale *Stufe* progressions such as Example 9—where the theory makes concrete predictions about perception, which, in my opinion, have a good chance of being borne out.

In Section 1, I argued that a theory might well be construed as both a theory of composition and a theory of perception—similar to a linguistic grammar. It may be noted that the aspects of ST that I have cited as most convincing with regard to compositional practice—namely, certain uses of *Stufen* and linear

71 Lerdahl (2001).

72 As Lerdahl and Jackendoff readily acknowledge (1983, 116, 273–74, 337–38), there are significant differences between prolongational reduction and Schenkerian reduction. Perhaps the biggest difference is that the elements of a prolongational reduction are verticalities, whereas a Schenkerian reduction may analyze the voices of a polyphonic texture as independent elements. Another significant difference is that in prolongational reduction, a V–I cadence is analyzed as a single indivisible event, in contrast to ST where the V and I are separate events.

73 Bigand and Parncutt (1999).

74 We should note that this mismatch between theory and perception regarding large-scale tonal structure also holds true for more conventional theoretical structures; see Note 58 above.

75 Lerdahl and Krumhansl (2007).

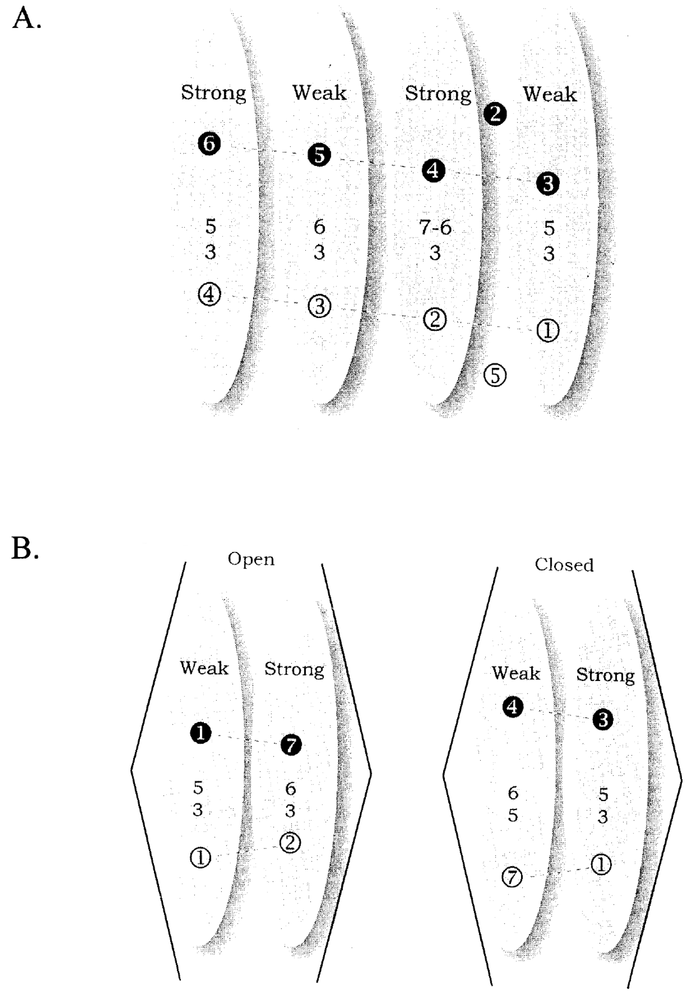
76 Lerdahl (2001).

progressions—are precisely those that I have argued for with regard to the modeling of perception. Perhaps these aspects of ST constitute a small part of the knowledge about common-practice music that composers and experienced listeners share—its "grammar," one might say—a body of knowledge also entailing principles of harmony, meter, phrase structure, melodic structure, form, and other things.

I readily acknowledge the subjective, speculative nature of the conclusions presented above. They are really just opinions at this point, and others may disagree with them. But I would argue that these opinions are given substance, and brought closer to *factual* claims, by the empirical methods I have proposed for testing them. That is to say: the claim "bass-line linear progressions are part of the compositional process" is very different from the claim "bass-line linear progressions, when combined with a finite-state model of surface harmonic progressions, can make harmonic progressions more normative—and thus predict them better—than a finite-state model alone." While neither of these claims has yet been tested in any systematic way, the latter claim is nearly, if not quite, a claim about observation terms, and thus, in principle, objectively testable.[77] With regard to perception, too, I have tried to go beyond mere intuitive guesses about what is perceptually valid and what is not, presenting concrete ways that such hypotheses might be tested.

If the opinions I have expressed above are correct, they suggest a new way of thinking about Schenkerian theory. Schenkerian structures, by this view, are not basic building blocks of tonal music. Rather, they are patterns that were used occasionally by tonal composers—perhaps quite often, but not all the time—and usually at fairly local levels. Using terms I have coined elsewhere,[78] they are "occasional" rather than "infrastructural" aspects of musical structure. A composer might, for example, generate the first and second themes of a sonata movement using traditional harmonic logic, but bring in a bass-line linear progression in the transition, and a large-scale *Stufe* progression at the end of the exposition. This view of *Stufen* and *Züge* brings to mind other similar "schemata" such as those posited by Leonard B. Meyer and Robert Gjerdingen—for example, the "$\hat{1}$–$\hat{7}$–$\hat{4}$–$\hat{3}$" pattern and the "Prinner" (see Example 17).[79] The *Zug* and the *Stufe* are of course somewhat more general than these, but they are similar in that they are compositional devices that tonal composers utilized here and there, from time to time, rather than omnipresent, essential components of tonal music.

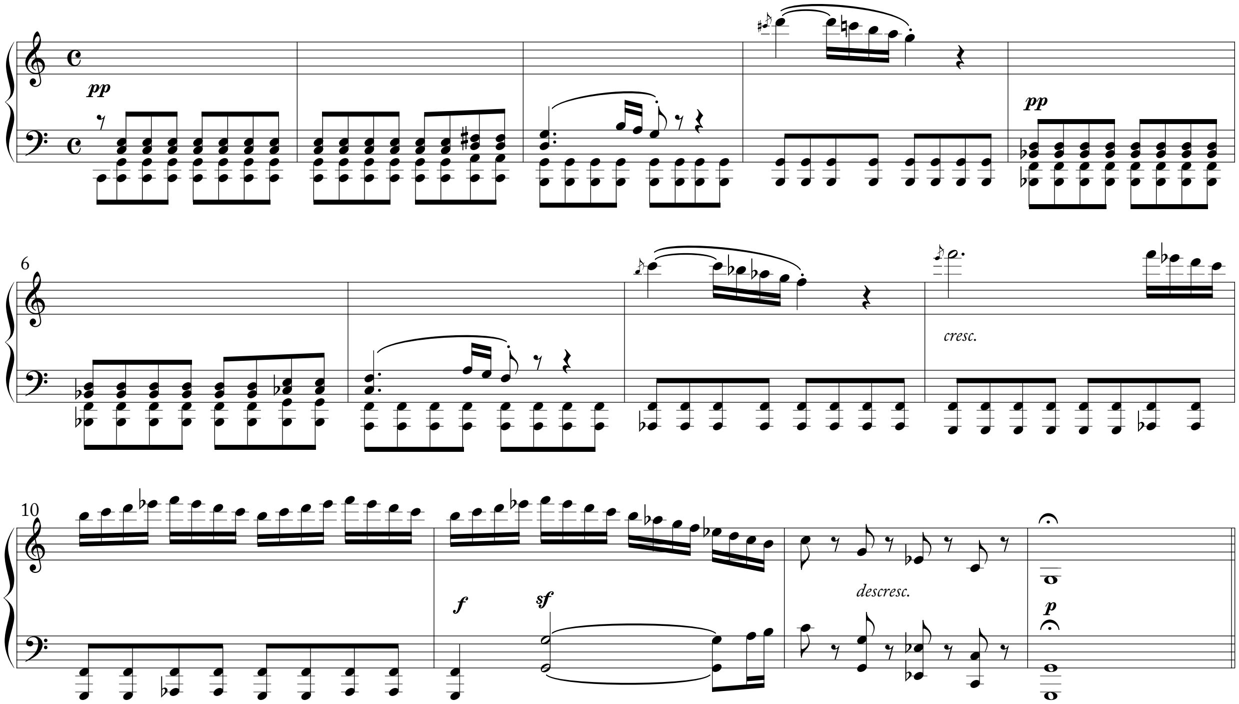
The view of Schenkerian structures as occasional schemata is also appealing from the standpoint of perception. I noted earlier that listeners have a remarkable ability to map one musical pattern on to another—to understand a complex structure as an elaboration of a simpler one. In a "theme and variations" situation such as Example 11, we compare segments that are unique



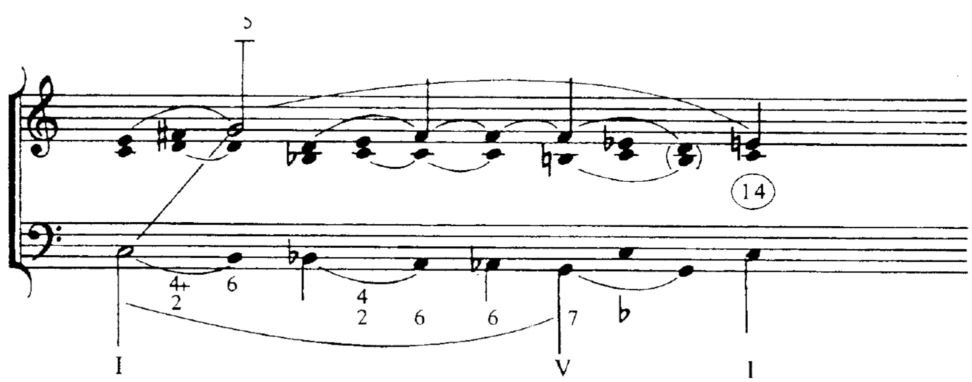
example 17. *Representations of two schemata of* Galant *music: the "Prinner" (a) and the "1–7–4–3" or "Meyer" (b) (From Gjerdingen [2007]. Reprinted by permission of Oxford University Press) Large ovals indicate the "events" of the schemata; black circles represent melodic scale degrees; open circles represent bass scale degrees; other numbers represent figured bass; "strong" and "weak" indicate the metrical placement of events*

to a particular piece. In the case of schemata, we compare a segment from a piece to a more abstract pattern stored in long-term memory. According to the view of schemata put forth by Gjerdingen,[80] a musical passage may evoke multiple schemata that overlap and even conflict (in terms of the notes that are identified as "structural"). And there is no requirement that a passage relate to any schema at all. Most of the structures posited in Schenkerian analyses, I would argue, are not recognized spontaneously, though they may seem plausible when juxtaposed with the musical passage under consideration. But in some cases, perhaps, the match between the schema and the passage is close enough that the schema is

77 As noted earlier, probabilistic methods could be used to quantify claims of this kind and test them rigorously. From a probabilistic viewpoint, to "make something more normative" means to assign it higher probability.
78 Temperley (2001, 3–4).
79 Meyer (1973) and Gjerdingen (2007).
80 Gjerdingen (2007).



EXAMPLE 18. *Beethoven, Piano Sonata No. 21 in C Major ("Waldstein"), Op. 53, I, mm. 1–13. Beethoven: Piano Sonata No. 21 in C Major, in Ludwig van Beethoven,* Sonaten für das Pianoforte, *3 vols., Ludwig van Beethovens Werke, Serie 16, Breitkopf und Härtel, Leipzig, 1862–90*



EXAMPLE 19. *From Beach (1987, 181): Analysis of Beethoven, Piano Sonata in C Major ("Waldstein"), Op. 53, I, mm. 1–13. © 1987, The Society for Music Theory, Inc. Used by permission. All rights reserved.*

evoked spontaneously; in such cases, the schema could truly be claimed to play a role in music perception.

As an example of what I propose, consider the opening of Beethoven's Sonata Op. 53 (the "Waldstein"), shown in Example 18. A Schenkerian analysis of this passage by David Beach[81] is shown in Example 19; the two other Schenkerian analyses I have found, by Gregory J. Marion and Lerdahl, are similar.[82] Some aspects of this passage seem to submit extremely well to a Schenkerian approach—most compellingly, the descending chromatic line in the bass. This, I would argue, is an outstanding example of a perceptually real (spontaneously heard) Schenkerian pattern, and also one that is quite plausible compositionally; as in Example 7, the linear bass pattern helps to explain (i.e., make more normative) the otherwise rather bizarre and unmotivated harmonic progression of the passage. The two ascending-third motions in the melody, E–F♯–G and D–E–F, are similarly convincing. The problem is what to do with mm. 9–13. Measure 9 presents a V^7 chord with a very prominent and

81 Beach (1987, 181).

82 Marion (1995, 12) and Lerdahl (2001, 206). Lerdahl calls his analysis "prolongational," but it is virtually indistinguishable from a Schenkerian analysis.

emphasized seventh (F) in the melody. From a conventional Schenkerian viewpoint, the prominence of this F seems to argue for it as a structural event, part of the middleground melodic line of the phrase; thus we expect it to make linear connections with events later in the phrase. But the most prominent notes of mm. 12 and 13 are C and G, respectively; there is no nice line here, even allowing for octave transfers. Beach's solution to this problem—advocated also by Marion and Lerdahl—is to posit a descending third F–E♭–D in mm. 9–13; the E♭ is (presumably) in the middle of the arpeggio in m. 12, and the D is not present at all but "implied."[83] Here, it seems to me, we are moving beyond anything that is plausible either perceptually or compositionally. Is it not possible that, at this point in the phrase, Beethoven simply was not thinking in terms of long-range linear patterns? Measures 9–13, I would suggest, are about other things: the move to the half-cadence (whether that move happens at m. 9 or at m. 13 I am not exactly sure); the strong dose of modal mixture (which has implications throughout the piece); the textural drama of two registrally extreme components coming together in the middle and then descending; the dramatic left-hand syncopation of m. 11; and the right-hand sixteenth-note motion, with its motivic connections to m. 4 and later passages. We certainly do not need to posit an underlying Schenkerian structure for mm. 9–13 to explain why this passage seems aurally satisfying, interesting, and logical in the context of the piece.

Again, I fully acknowledge the speculative nature of this essay and the conclusions I have drawn. Ultimately, quantitative tests—corpus analyses and music perception experiments—will be necessary to evaluate them. Undoubtedly, too, there are other ways of evaluating the theory beyond those discussed here. I have offered some suggestions as to how ST can make testable predictions about tonal music, and how it might yield substantive predictions about music perception; perhaps others can find additional examples, demonstrating the predictive power of aspects of ST that I have not considered here. The fundamental questions are simple ones: In what ways does ST predict what happens, and does not happen, in tonal music? And in what ways can it be shown to predict phenomena of music perception, and thus potentially explain or illuminate aspects of musical experience? These are not the only ways that ST might be evaluated; and even if it failed completely in these respects, we might still judge the theory to have value in other ways. But explaining the mental representations involved in the composition and perception of tonal music is surely among the valid goals of our discipline; it seems reasonable to inquire whether Schenkerian theory can contribute to these goals.

83 Assuming the F in m. 9 is structural, a question arises as to whether it is prolonged through mm. 9–13 to resolve on the E in m. 14 (as in Beach [1987, 181] and Lerdahl [2001, 206]), or whether it resolves on the E♭ in m. 12, with the implied D of m. 13 being the structural ending of the phrase (as in Marion [1995, 12]); but this question is not vital for our purposes.

WORKS CITED

Aldwell, Edward, and Carl Schachter. 2003. *Harmony and Voice Leading*. 3rd ed. Belmont [CA]: Thomson/Schirmer.

Ayotte, Benjamin McKay. 2004. *Heinrich Schenker: A Guide to Research*. New York: Routledge.

Babbitt, Milton. 1952. Review of *Structural Hearing* by Felix Salzer. *Journal of the American Musicological Society* 5 (3): 260–65.

Beach, David. 1983. "Schenker's Theories: A Pedagogical View." In *Aspects of Schenkerian Theory*. Ed. David Beach. 1–38. New Haven: Yale University Press.

———. 1987. "On Analysis, Beethoven, and Extravagance: A Response to Charles J. Smith." *Music Theory Spectrum* 9: 173–85.

Benjamin, William E. 1981. "Schenker's Theory and the Future of Music." *Journal of Music Theory* 25 (1): 155–73.

Berry, David Carson. 2004. *A Topical Guide to Schenkerian Literature: An Annotated Bibliography with Indices*. Hillsdale [NY]: Pendragon Press.

Bigand, Emmanuel, and Richard Parncutt. 1999. "Perceiving Musical Tension in Long Chord Sequences." *Psychological Research* 62 (4): 237–54.

Brown, Matthew G. 1989. "A Rational Reconstruction of Schenkerian Theory." Ph.D. diss., Cornell University.

Brown, Matthew. 2005. *Explaining Tonality: Schenkerian Theory and Beyond*. Rochester: University of Rochester Press.

Brown, Matthew, Douglas Dempster, and Dave Headlam. 1997. "The ♯IV(♭V) Hypothesis: Testing the Limits of Schenker's Theory of Tonality." *Music Theory Spectrum* 19 (2): 155–83.

Cadwallader, Allen Clayton, and David Gagné. 2007. *Analysis of Tonal Music: A Schenkerian Approach*. 2nd ed. New York: Oxford University Press.

Chomsky, Noam. 1957. *Syntactic Structures*. The Hague: Mouton.

Conklin, Darrell, and Ian Witten. 1995. "Multiple Viewpoint Systems for Music Prediction." *Journal of New Music Research* 24 (1): 51–73.

Cook, Nicholas. 1987a. *A Guide to Musical Analysis*. New York: W. W. Norton.

———. 1987b. "The Perception of Large-Scale Tonal Closure." *Music Perception* 5 (2): 197–206.

Covach, John, and Graeme M. Boone. 1997. *Understanding Rock: Essays in Musical Analysis*. Oxford: Oxford University Press.

Cuddy, Lola L., and Carol A. Lunney. 1995. "Expectancies Generated by Melodic Intervals: Perceptual Judgments of Melodic Continuity." *Perception and Psychophysics* 57 (4): 451–62.

Dibben, Nicola. 1994. "The Cognitive Reality of Hierarchic Structure in Tonal and Atonal Music." *Music Perception* 12 (1): 1–25.

Fodor, Jerry. 1984. "Observation Reconsidered." *Philosophy of Science* 51 (1): 23–43.

Forte, Allen. 1977. "Schenker's Conception of Musical Structure." In *Readings in Schenker Analysis and Other Approaches*. Ed. Maury Yeston. 3–37. New Haven: Yale University Press.

Forte, Allen, and Steven Gilbert. 1982. *An Introduction to Schenkerian Analysis*. New York: W. W. Norton.

Fux, Johann Joseph. [1725] 1971. *The Study of Counterpoint from Johann Joseph Fux's* "Gradus ad Parnassum." Trans. and ed. Alfred Mann. New York: W. W. Norton. Originally published as *Steps to Parnassus: The Study of Counterpoint*.

Garman, Michael. 1990. *Psycholinguistics*. Cambridge: Cambridge University Press.

Gauldin, Robert. 2004. *Harmonic Practice in Tonal Music*. 2nd ed. New York: W. W. Norton.

Gilbert, Édouard, and Darrell Conklin. 2007. "A Probabilistic Context-Free Grammar for Melodic Reduction." Paper presented at the International Workshop on Artificial Intelligence and Music, IJCAI-07, 6–12 January, Hyderabad, India.

Gjerdingen, Robert O. 2007. *Music in the Galant Style*: *Being an Essay on Various Schemata Characteristic of Eighteenth-Century Music for Courtly Chambers, Chapels, and Theaters, Including Tasteful Passages of Music Drawn from Most Excellent Chapel Masters in the Employ of Noble and Noteworthy Personages, Said Music All Collected for the Reader's Delectation on the World Wide Web*. New York: Oxford University Press.

Huron, David. 2001. "Tone and Voice: A Derivation of the Rules of Voice-Leading from Perceptual Principles." *Music Perception* 19: 1–64.

———. 2006. *Sweet Anticipation: Music and the Psychology of Expectation.* Cambridge [MA]: The MIT Press.

Judd, Cristle Collins. 1992. "Modal Types and *Ut, Re, Mi* Tonalities: Tonal Coherence in Sacred Vocal Polyphony from About 1500." *Journal of the American Musicological Society* 45 (3): 428–67.

Kassler, Michael. 1977. "Explication of the Middleground of Schenker's Theory of Tonality." *Miscellanea Musicologica* 9: 72–81.

Keiler, Allan. 1978. "The Empiricist Illusion: Narmour's *Beyond Schenkerism*." *Perspectives of New Music* 17 (1): 161–95.

Kerman, Joseph. 1985. *Contemplating Music: Challenges to Musicology*. Cambridge [MA]: Harvard University Press.

Kostka, Stefan. 1995. *Instructor's Manual to Accompany "Tonal Harmony."* New York: McGraw-Hill.

Kostka, Stefan, and Dorothy Payne. 1995. *Workbook for "Tonal Harmony."* New York: McGraw-Hill.

Krumhansl, Carol L. 1990. *Cognitive Foundations of Musical Pitch*. New York: Oxford University Press.

Laitz, Steven G. 2008. *The Complete Musician: An Integrated Approach to Tonal Theory, Analysis, and Listening.* 2nd ed. New York: Oxford University Press.

Larson, Steve. 1997. "The Problem of Prolongation in *Tonal* Music: Terminology, Perception, and Expressive Meaning." *Journal of Music Theory* 41: 101–36.

———. 1997–98. "Musical Forces and Melodic Patterns." *Theory and Practice* 22–23: 55–71.

———. 1998. "Schenkerian Analysis of Modern Jazz: Questions about Method." *Music Theory Spectrum* 20 (2): 209–41.

———. 2004. "Musical Forces and Melodic Expectations: Comparing Computer Models and Experimental Results." *Music Perception* 21 (4): 457–98.

Lerdahl, Fred. 2001. *Tonal Pitch Space*. New York: Oxford University Press.

Lerdahl, Fred, and Carol L. Krumhansl. 2007. "Modeling Tonal Tension." *Music Perception* 24 (4): 329–66.

Lerdahl, Fred, and Ray Jackendoff. 1983. *A Generative Theory of Tonal Music*. Cambridge [MA]: The MIT Press.

Maisel, Arthur. 1999. "Voice Leading as Drama in 'Wozzeck.' " In *Schenker Studies 2*. Ed. Carl Schachter and Hedi Siegel. 160–91. Cambridge: Cambridge University Press.

Manning, Christopher D., and Hinrich Schütze. 1999. *Foundations of Statistical Natural Language Processing*. Cambridge [MA]: The MIT Press.

Marion, Gregory J. 1995. "Inciting Transformational Insights." *Intégral* 9: 1–32.

Marsden, Alan. 2010. "Schenkerian Analysis by Computer: A Proof of Concept." *Journal of New Music Research* 3: 269–89.

Mavromatis, Panayotis. 2005. "A Hidden Markov Model of Melody Production in Greek Church Chant." *Computing in Musicology* 14: 93–112.

Meyer, Leonard B. 1973. *Explaining Music: Essays and Explorations*. Berkeley: University of California Press.

Narmour, Eugene. 1977. *Beyond Schenkerism: Toward an Alternative in Music Analysis.* Chicago: The University of Chicago Press.

———. 1990. *The Analysis and Cognition of Basic Melodic Structures: The Implication-Realization Model*. Chicago: The University of Chicago Press.

Nattiez, Jean-Jacques. 1990. *Music and Discourse: Toward a Semiology of Music.* Trans. Carolyn Abbate. Princeton: Princeton University Press.

Neumeyer, David, and Susan Tepping. 1992. *A Guide to Schenkerian Analysis*. Englewood Cliffs [NJ]: Prentice-Hall.

Palmer, Carol, and Carol L. Krumhansl. 1990. "Mental Representations for Musical Meter." *Journal of Experimental Psychology: Human Perception and Performance* 16 (4): 728–41.

Peel, John, and Wayne Slawson. 1984. Review of *A Generative Theory of Tonal Music* by Fred Lerdahl and Ray Jackendoff. *Journal of Music Theory* 28 (2): 271–94.

Roig-Francoli, Miguel A. 2003. *Harmony in Context*. Boston: McGraw-Hill.

Rosen, Charles. 1971. *The Classical Style: Haydn, Mozart, Beethoven*. New York: Viking Press.

Rothgeb, John. 1975. "Strict Counterpoint and Tonal Theory." *Journal of Music Theory* 19 (2): 260–84.

Salzer, Felix. 1952. *Structural Hearing: Tonal Coherence in Music*. New York: Dover Publications.

Schachter, Carl. 1976. "Rhythm and Linear Analysis: A Preliminary Study." *The Music Forum* 4: 281–334. Reprinted in *Unfoldings: Essays in Schenkerian Theory and Analysis.* Ed. Joseph N. Straus. 17–53. New York: Oxford University Press, 1999.

———. 1981. "A Commentary on Schenker's *Free Composition*." *Journal of Music Theory* 25 (1): 115–42. Reprinted in *Unfoldings: Essays in Schenkerian Theory and Analysis.* Ed. Joseph N. Straus. 184–208. New York: Oxford University Press, 1999.

———. 1987. "Analysis by Key: Another Look at Modulation." *Music Analysis* 6 (3): 289–318. Reprinted in *Unfoldings: Essays in Schenkerian Theory and Analysis.* Ed. Joseph N. Straus. 134–60. New York: Oxford University Press, 1999.

Schellenberg, E. Glenn. 1996. "Expectancy in Melody: Tests of the Implication-Realization Model." *Cognition* 58 (1): 75–125.

Schenker, Heinrich. [1924] 2004. *Der Tonwille*. Trans. Ian Bent, William Drabkin, Joseph Dubiel, Timothy Jackson, Joseph Lubben, and Robert Snarrenberg. Ed. William Drabkin. New York: Oxford University Press.

———. [1930] 1994. *The Masterwork in Music* (*Das Meisterwerk in der Musik: Ein Jahrbuch*). Vol. 3. Trans. Ian Bent, William Drabkin, Richard Kramer, John Rothgeb, and Hedi Siegel. Ed. William Drabkin. Cambridge: Cambridge University Press.

———. [1933] 1969. *Five Graphic Music Analyses* (*Fünf Urlinie-Tafeln*). New York: Dover Publications.

———. [1935] 1979. *Free Composition* (*Der freie Satz*). Trans. and ed. Ernst Oster. New York: Longman.

Schmuckler, Mark A. 1989. "Expectation and Music: Investigation of Melodic and Harmonic Processes." *Music Perception* 7 (2): 109–50.

Schoenberg, Arnold. [1954] 1969. *Structural Functions of Harmony*. Rev. ed. Ed. Leonard Stein. New York: W. W. Norton.

Serafine, Mary Louise, Noah Glassman, and Cornell Overbeeke. 1989. "The Cognitive Reality of Hierarchic Structure in Music." *Music Perception* 6 (4): 397–430.

Smith, Charles J. 1986. "The Functional Extravagance of Chromatic Chords." *Music Theory Spectrum* 8: 94–139.

Temperley, David. 1999. "The Question of Purpose in Music Theory: Description, Suggestion, and Explanation." *Current Musicology* 66: 66–85.

———. 2001. *The Cognition of Basic Musical Structures*. Cambridge [MA]: The MIT Press.

———. 2007. *Music and Probability*. Cambridge [MA]: The MIT Press.

———. 2009. "A Statistical Analysis of Tonal Harmony." http://www.theory.esm.rochester.edu/temperley/kp-stats (accessed 30 March 2010).

———. 2010. "Modeling Common-Practice Rhythm." *Music Perception* 27 (5): 355–76.

Von Hippel, Paul. 2000. "Redefining Pitch Proximity: Tessitura and Mobility as Constraints on Melodic Intervals." *Music Perception* 17 (3): 315–27.

Von Hippel, Paul, and David Huron. 2000. "Why Do Skips Precede Reversals? The Effect of Tessitura on Melodic Structure." *Music Perception* 18 (1): 59–85.

Music Theory Spectrum, Vol. 33, Issue 2, pp. 146–168, ISSN 0195-6167, electronic ISSN 1533-8339. DOI: 10.1525/mts.2011.33.2.146