SIBELIUS ACADEMY
Department of Composition and Music Theory
Studia Musica 23, 2004

Väisälä, Olli
Prolongation in Early Post-Tonal Music: Analytical Examples and Theoretical Principles

---

ABSTRACT

The studies that compose this dissertation analyze a selection of pieces of early post-tonal music (by Debussy, Scriabin, Schoenberg, Berg, and Webern) on the basis of the notion of prolongation. They also discuss extensively the theoretical principles of post-tonal prolongation and, to some extent, the relationships of these principles with psychoacoustical phenomena.

Prolongation is a key notion in Schenkerian analysis of conventionally tonal music, and there have been various attempts to generalize this notion to meet the demands of post-tonal music. However, whereas conventional Schenkerian analysis is regulated by well-defined theoretical principles related to the normative referential position of the triad, purported prolongational analyses of post-tonal music have, in general, remained unsatisfactory, owing to the lack of comparable theoretical principles. The present studies determine such principles for the selection of works analyzed, on the basis of non-triadic referential harmonies.

The theoretical discussion draws on Joseph Straus's (1987) four conditions for prolongation, a well-known formulation of pitch-based functional norms required by prolongation. However, the approach differs from Straus's in its conception of harmonies and intervals, by incorporating aspects outside the purview of pitch-class set theory; it turns out that this decisively improves the prospects for post-tonal prolongation. Two such aspects are discussed. The first is registration; it is argued that registral distinctions (such as between certain complementary intervals) are crucial for functional distinctions in almost any kind of prolongational organization. The second—which pertains to a more limited repertoire—is rootedness, a property stemming from approximate correspondences between musical intervals and those in the harmonic series. Theoretical principles, such as these two aspects, are considered from two angles: how they illuminate the works analyzed, and how they relate with perceptual (psychoacoustical) principles.

In the present selection of compositions, the theoretical foundation enables prolongational analyses whose descriptive power is largely comparable to that of conventional Schenkerian analyses. While several of the theoretical principles are likely to have general significance for the illumination of musical organization in comparable repertoire, only further studies can decide the extent to which this illumination actually amounts to the revelation of prolongational structures.

TIIVISTELMÄ

Tämä väitöskirja muodostuu tutkimuksista, joissa analysoidaan valikoima varhaista post-tonaalista musiikkia (Debussyn, Skrjabinin, Schönbergin, Bergin ja Webernin teoksia) prolongaation käsitteen pohjalta. Niissä myös tarkastellaan laajasti post-tonaalisen prolongaation teoreettisia periaatteita ja jossain määrin näiden periaatteiden suhteita psykoakustisiin ilmiöihin.

Prolongaatio on keskeinen käsite perinteisen tonaalisen musiikin Schenker-analyysissä, ja on tehty monia yrityksiä käsitteen yleistämiseksi post-tonaaliseen musiikkiin soveltuvaksi. Siinä missä konventionaalista Schenker-analyysiä ohjaavat kolmisoinnun referentiaaliseen asemaan liittyvät hyvin määritellyt teoreettiset periaatteet, esitetyt post-tonaalisen musiikin prolongaatioanalyysit ovat kuitenkin yleisesti ottaen jääneet epätyydyttäviksi, koska vastaavat teoreettiset periaatteet ovat puuttuneet. Tässä tutkimuksessa määritetään sellaiset periaatteet analysoiduille teoksille ei-kolmisointuisiin referenssiharmonioihin pohjautuen.

Teoreettisten kysymysten käsittely pohjautuu Joseph Strausin (1987) neljään prolongaation ehtoon, tunnettuun muotoiluun prolongaation edellyttämistä säveltasoon perustuvista funktionaalisista normeista. Käsittely poikkeaa kuitenkin Strausista harmonia- ja intervallikäsityksen osalta ottamalla huomioon säveljoukkoluokkien teorian ulkopuolisia tekijöitä; osoittautuu, että tämä ratkaisevasti parantaa post-tonaalisten prolongaatiorakenteiden esiintymis- ja löytymismahdollisuuksia. Tällaisia tekijöitä on kaksi. Ensinnäkin rekisteriasettelu: on perusteita olettaa, että rekisteriin perustuvat erotukset (esimerkiksi tiettyjen käännösintervallien väliset) vaikuttavat ratkaisevasti funktionaalisiin erotuksiin melkeinpä millaisessa tahansa prolongaatioorganisaatiossa. Toinen tekijä – jolla on merkitystä rajoitetummalle ohjelmistolle - on pohjasävelisyys, ominaisuus, joka perustuu likiarvoisiin vastaavuuksiin musiikillisten ja yläsävelsarjassa esiintyvien intervallien välillä. Teoreettisia periaatteita, kuten edellä mainittuja tekijöitä, tarkastellaan kahdelta kannalta: miten ne valaisevat analysoituja tekijöitä ja miten ne suhteutuvat havaintoa koskeviin (psykoakustisiin) periaatteisiin.

Tutkitussa teosvalikoimassa teoreettinen perusta tekee mahdolliseksi prolongaatioanalyysit, joiden kuvausvoima on pitkälti verrattavissa konventionaalisiin Schenker-analyyseihin. Vaikka useilla teoreettisista periaatteista lienee yleistä merkitystä musiikillisen organisaation valaisijana vastaavassa ohjelmistossa, vain lisätutkimukset voivat selvittää, missä määrin tämä valaistus riittää paljastamaan varsinaisia prolongaatiorakenteita.

ORIGINAL PUBLICATIONS

This thesis consists of an introductory essay and the following three articles. In the introductory essay, these articles will be referred to by boldface Roman numerals, as indicated. Articles **I** and **II** are attached at the end of this dissertation as photocopies. Article **III** is attached as a manuscript.

I Väisälä, Olli. 1999. "Concepts of Harmony and Prolongation in Schoenberg's Op. 19/2." *Music Theory Spectrum* 21: 230–59.

II Väisälä, Olli. 2002. "Prolongation of Harmonies Related to the Harmonic Series in Early Post-Tonal Music." *Journal of Music Theory* 46: 207–83.

III Väisälä, Olli. 2006. "New Theories and Fantasies on the Music of Debussy: Post-Triadic Prolongation in *Ce qu'a vu le vent d'ouest* and Other Examples." In *Selected Essays from the Third International Schenker Symposium*. Hildesheim: Georg Olms, 2006.

[*Note*. The electronic version of this dissertation does not contain attachments of articles **I**–**III**. The information concerning article **III**, which was published after the printed version of the dissertation (2004), is updated above.]

ACKNOWLEDGEMENTS

I would like to express my thanks to the supervisors of this thesis. Prof. Lauri Suurpää offered numerous detailed and valuable suggestions concerning both theoretical and analytical issues. Prof. Ilkka Oramo's attitude towards my theoretical endeavors has always been most encouraging.

Several other teachers and colleagues have influenced my musical understanding and thus this thesis. Risto Väisänen deserves a very special mention. Without his *satsioppi* (harmony and counterpoint) tuition few of my present analytical and theoretical ideas would probably ever have arisen.

Prof. Joseph Straus offered helpful comments on a preliminary version of the analysis of Debussy's *Voiles* (later incorporated in **II**), triggering, in a sense, the process which eventually led to this thesis. Prof. Richard Parncutt shared his psychoacoustical expertise via e-mail correspondence, commenting parts of **I** and **II** in detail.

My English in the introductory essay was checked by Timothy Page. In the articles, and manuscripts thereof, language editing was done by Susan Sinisalo, Phillip Money, and Andrew Bentley. Mr. Bentley also helped me with electronical equipment in arranging informal experiments in virtual-pitch perception.

The editors of *Music Theory Spectrum*, *Journal of Music Theory*, and the forthcoming *Third International Schenker Symposium* publication, Philip Lambert, Eric Drott, David Clampitt, and Allen Cadwallader have also influenced the formulation of the texts, in some cases quite considerably.

Prof. Joel Lester, who reviewed this thesis alongside Straus, made valuable suggestions for emendations. Unfortunately, for reasons of schedule, I could not follow what was perhaps the most important of them: a call for more extended considerations on the scope of the applicability of the present approach. Clarification of this issue has to wait for future research (by anyone interested).

Finally, I thank my wife Soile for true companionship during times of toil and times of enjoyment.

COPYRIGHT ACKNOWLEDGEMENTS

TABLE OF CONTENTS FOR THE INTRODUCTORY ESSAY

Introductory Essay

1. INTRODUCTION

The studies that compose this dissertation have two main aims. The first is to offer a new kind of analytical illumination for a selection of stylistically innovative (post-tonal)[1] compositions from the early 20th century. The second is to discuss and develop theoretical principles relevant to such analysis, focusing on the notion of post-tonal prolongation. The discussion of theoretical principles considers, apart from their productivity in analysis, their relationships with certain psychoacoustical phenomena. (The relationships between analysis, music theory, and psychoacoustics are addressed in section 2 below.)

The analyzed repertoire comprises the following works: in **I**, Arnold Schoenberg, piano pieces op. 19 no. 2 (1911) and op. 11 no. 2 (1909); in **II**, Alexander Scriabin, *Vers la flamme* op. 72 (1914), Alban Berg, song op. 2 no. 2 ("Schlafen trägt man mich") (1909–10), Claude Debussy, *Voiles* (first book of *Préludes* for piano, no. 2) (1909), Anton Webern, song. op. 3 no. 1 ("Dies ist ein Lied") (1909–10); in **III**, Debussy, *Ce qu'a vu le vent d'ouest* (first book of *Préludes*, no. 7) (1910) and parts of other works.[2]

The basis of the present theoretical discussion, the notion of prolongation, has its origins in Schenkerian analysis of conventional tonal music. Because of the productivity of this notion in the analysis of tonal music—concerning both small and large temporal spans—several scholars have attempted to generalize it to meet the demands of post-tonal music. One approach to this, adopted in the present studies, is based on the idea that, whereas Schenkerian analysis views conventional music as developing from the prolongation (or "composing-out")[3] of the major or minor triad, in post-tonal music other, more or less work-specific referential harmonies replace the triad as objects of prolongation.[4]

---

[1] In the present studies, "post-tonal" refers to any kind of 20th-century music whose organization does not adhere to the organizational norms of conventional or "common practice" tonality, which manifest the governing position of the major or minor triad. Being post-tonal in the present use of the word does not rule out the significance of non-conventional use of tonal centers, or, for that matter, allusions to conventional tonality. I have used the word "post-tonal" rather than "atonal" in order to avoid the impression that the music discussed is strongly antithetical to "tonal."

[2] Most of the analyses cover entire works in a detailed fashion. However, for Schoenberg's op. 11/2 and Scriabin's *Vers la flamme*, only the opening is analyzed in a detailed way, whereas the overall organization is treated more cursorily.

[3] On these concepts see section 3, especially note 42.

[4] Examples of the "referential harmony" approach include Travis 1959 and 1966, some analyses in Morgan 1976, and Laufer 1991 (and several other unpublished analyses by Laufer). This approach is also adumbrated in some analyses in Katz (1945/1972), especially in that of the opening of Debussy's *Voiles*, which is essentially similar to the present one (ibid.: Example 93; cf. **II**: Examples 15–16).

Another approach has been to abide more closely by original Schenkerian notions, enriching them with some extensions and additions, such as Felix Salzer's (1952/1982) "contrapuntal-structural tones," "'independent' voice leading," and "color chords." In general, such analysis has been most successful when the musical style is, in accordance with the analytical approach, not too far removed from conventional tonality. However, some scholars, e.g., Baker (1990) and Cinnamon (1993), have also applied Schenkerian concepts, such as *Ursatz*, in a more or less direct way to pieces whose foreground harmonic vocabulary and syntax bear little evidence of the norms of conventional tonality and the referential status of the triad, on which the justification of the *Ursatz* ultimately relies. This approach contrasts strongly with the present one, in which substantiation for structural principles is sought from the material actually present in the music. (See especially **I**: 254–59, in which I compare Cinnamon's readings of Schoenberg's op. 11/2 with those achieved by the present approach.).

Several authors have also found the concept of post-tonal prolongation to be problematic (Oster 1960, Laufer 1981 [161], Baker 1983, Strauss 1987).[5] In tonal music, prolongational analysis is regulated by well-known pitch-based norms of syntax (of consonance and dissonance, for example) and scholars have found it questionable whether similar or comparable norms may be traced or postulated for post-tonal music. Most often, post-tonal prolongational analyses have failed to explicate the structural principles on which they are based. For whatever merits such analyses may have had in illustrating significant musical intuitions, the nature of such intuitions remains unclear owing to defective theoretical groundwork.[6] Therefore, the suggested interpretations of post-tonal prolongation have been criticized for being "somewhat arbitrarily based" (Baker 1983: 168) or not genuinely prolongational (Straus 1987).

The most specific explication of the distinctive characteristics of prolongational structures and their applicability to post-tonal music is Strauss 1987. Straus postulates four conditions for prolongation, which call for the existence of certain pitch-based norms relevant to structural relationships. In tonal music, these conditions are met in a way that largely stems from the normative status of the triad. However, there is nothing inherent in the conditions that would make it impossible for them to be met in ways that differ from those in conventional, triadic tonality (ibid.: 4, 7). Nevertheless, through his studies of a few analytical examples, Strauss concludes that this is not the case in "the most significant post-tonal music."[7] Straus goes on

---

[5] Interestingly, both Baker and Laufer have, since writing these studies, contributed to the literature of quasi-Schenkerian analysis of post-tonal music, perhaps suggesting that despite the difficulty to formulate theoretical principles for such analysis the approach has intuitively appealing features.

[6] Some exceptions to this should be acknowledged. Lester 1970 includes rather thorough discussion of the theoretical principles underlying his analyses of Schoenberg's *Serenade* op. 24. Lester's discussion agrees in several points with the present considerations on the *proximity principle* (section 3.2.3.3; **I**: 235 ff.; **II**: section 1.3). More recent work on explicating principles for some kind of post-tonal or atonal "prolongation" has been undertaken by Fred Lerdahl (1989, 1999, 2001: chapters 6–8). An important common feature between Lerdahl's work (1999 and 2001) and mine is the concern for underlying psychoacoustical factors (such as virtual pitch, roughness, and streaming), although the approaches are otherwise different (see notes 7 and 84 below).

[7] In recent years, relatively little work has been written on post-tonal prolongation, suggesting that most theorists have agreed with Straus's conclusions. In fact, Straus (1987: 1) already notes that "[w]ith a few

to suggest that large-scale analysis of post-tonal music should, in general, be based on the principle of *association* rather than prolongation (ibid.: 8 ff.).[8]

The present studies adopt Strauss's conditions as the basis of the notion of prolongation, but come to partly different conclusions as to their applicability in post-tonal music. The main reason for this difference lies in a different conception of harmonies and intervals. Straus identifies harmonies as pitch-class sets and intervals as interval classes, which involves the presupposition of full octave equivalence or unrestricted registral freedom. In the present studies, octave equivalence is restricted in variable degrees; in other words, registral disposition is regarded as essential for the identity of harmonies and intervals. For example, in **I**, I suggest that some of Schoenberg's music treats the major seventh (or registrally ordered interval 11; see section 4.1.1) as a functional consonance and the minor second/ninth (or registrally ordered interval 1) as a functional dissonance, even though they both represent interval class 1. Since the first of Straus's conditions concerns consonance and dissonance, the recognition of such distinctions have fundamental consequences for considerations of prolongation. This becomes especially evident in considerations of Schoenberg's op. 19/2, an example treated in both Straus 1987 and **I**.

Apart from registral distinctions, the present studies also bring up another aspect of harmonies and intervals that is not allowed for by set-theoretical concepts. This aspect is *rootedness* and derives from the (approximate) correspondence between musical intervals and

---

exceptions, theorists have virtually ceased to produce prolongational analyses of post-tonal music," a tendency which Straus's article has further promoted. Paul Wilson (1992: 42), for example, goes so far as to assert, drawing on Baker 1983 and Straus 1987, that "[A]ny attempt to find complete and convincing analogies to prolongation in post-tonal music is doomed to failure." Only few analysts (Morrison 1991, Pearsall 1991) have suggested that some of Straus's conditions are fulfilled in non-conventional ways in post-tonal repertoire. On the other hand, others—e.g., Lerdahl 1989, Travis 1990—have contested the validity of Straus's conditions, charging him with circular reasoning (see section 3.2.2). Lerdahl 1989 argues that atonal prolongation is based on "salience conditions" rather than the "stability conditions" required by Straus 1987 (although Lerdahl 1999 and 2001 re-introduce some pitch-based stability factors). For this reason, Lerdahl's approach deviates clearly from the present one, which is extensively based on Straus's conditions. It should also be observed that Lerdahl's approach, unlike the present one, is not actually based on the generalization of Schenkerian theory, but of Lerdahl and Jackendoff's (1983) version of prolongational theory (for a brief comparison of these theories, see, for example, Lerdahl 2001: 10). **I** (251–2) illustrates the difference between the results given by the approaches by comparing readings of Schoenberg's op. 19/2.

[8] More recently, Straus (1997a and 2003) has advocated another, *transformational* approach to "voice leading in atonal music." This approach is based on transposition and inversion operations between harmonies, interpreted as pitch-class sets. For a succession of two sets related by such operations, the same operations determine the "voice leading"; for example, for two subsequent sets related by T_4, all "voices" move four semitones up in pitch-class space. (Straus 2003 discusses and defines "fuzzy" extensions of transposition and inversion, enabling any two sets to be related by such operations.) While Straus calls such imaginary motion "voice leading"—viewing it as comparable with voice leading in conventional tonality and in prolongational and associational models of post-tonal music—it actually contrasts sharply with what is normally understood by voice leading. Firstly, Straus's "transformational voice leading" is determined by the choice of harmonies (pitch-class sets), whereas voice leading in the traditional sense (or in the prolongational and associational models of post-tonal music) is an independent aspect of organization: there are several ways in which voices may be led in a progression of harmonies. Secondly, both prolongational and associational voice leading are strongly related with supporting perceptual factors, which is not the case for "transformational voice leading." While Straus is certainly justified in distinguishing between prolongation and association, these factors often coincide to a significant degree, whereas "transformational voice leading" would seem to be an altogether different aspect of organization.

those in the harmonic series. (The psychoacoustical underpinnings of this concept, based on *virtual pitch*, are discussed in section 5.2.1). Whereas registral distinctions are relevant to the conditions of prolongation in all the examples, the relevance of rootedness is most evident in a more limited repertoire, discussed in **II** and **III**.

The significance of both registration and rootedness is considered from two angles in the present studies: with respect to analytical consequences, on the one hand, and to perceptual relevance, on the other. Perceptual considerations are made partly on an informal (subjective) basis and partly by discussing specific connections between principles of musical organization and psychoacoustics.

The present analytical examples have been purposely chosen because of the clarity in which they demonstrate the possibility of post-tonal prolongation. While all of them show the conditions of prolongation to be clearly operative, one should be cautious about drawing conclusions that apply generally to post-tonal music. The application of similar principles to the analysis of other comparable repertoire, even to music by the same composers, does not necessary reveal prolongational organization of comparable clarity. Two more modest conclusions seem justified, however. First, *insofar* as prolongational organization is to be found in comparable repertoire, features in the present approach—in particular, register-sensitivity—are likely to be relevant to it. Second, the present prolongational considerations demonstrate the organizational potential of features that lie outside the standard set-theoretical conception of harmony; awareness of such potential may be important for analysis of post-tonal music even when the organization does not fulfill the conditions of prolongation as pervasively as the present examples.

In addition to post-tonal music that meets the conditions of prolongation only partially or temporarily, there is another borderline area that may be illuminated by aspects of the present approach: the large repertoire of music near the borderline between tonal and post-tonal, whose illumination may require a double perspective. In the present studies, such a double perspective is evident in the discussion of Berg's op. 2 in **II** and in some Debussy analyses in **III** (see especially the concluding discussion of *L'Isle joyeuse*).

The present essay serves as an introduction to the three articles. It contains a unified presentation of relevant theoretical and psychoacoustical issues (whereas the theoretical discussion in each of the articles concentrates on topics relevant to the respective analytical considerations). In addition, I have seized the opportunity to elaborate on some special issues that are only briefly touched upon in the articles. By contrast, the analytical observations, which occupy a major position in the articles, are discussed only for the purpose of illustrating the theoretical discussion.

Section 2 discusses the relationships between analysis, theory, and psychoacoustics as they pertain to the present studies. Section 3 focuses on the concept of prolongation, discussing the characteristic features of prolongation on the basis of conventional tonality, relating

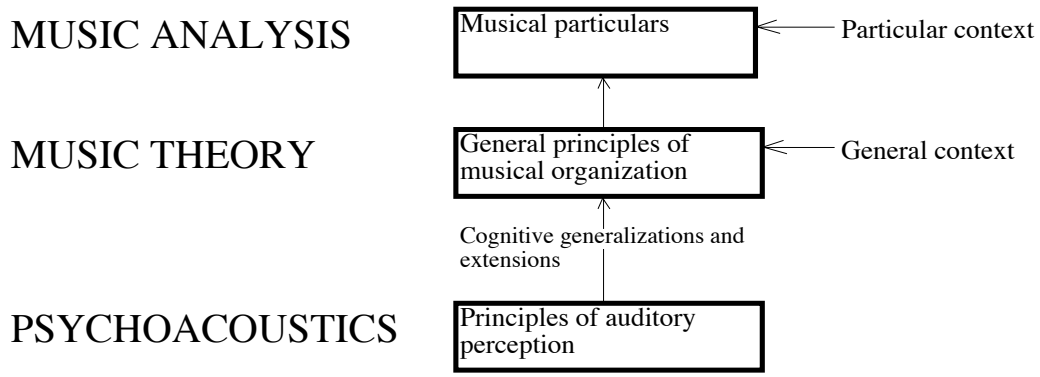
prolongation to other aspects of organization, and discussing Straus's conditions in detail. Section 4 discusses aspects of harmonic and intervallic concepts, and section 5 discusses psychoacoustical phenomena relevant to the present theoretical principles. Section 6 considers the concept of referential harmony, summarizes the most important principles relevant to the conditions of prolongation, and discusses briefly the conclusions that the present results imply for the analysis of post-tonal music.

2. ANALYSIS, THEORY, AND PSYCHOACOUSTICS

2.1 A MODEL OF RELATIONSHIPS

Example 1 presents a model of the relationships between analytical, theoretical, and psychoacoustical considerations, as evident in the analytical explanations of the present studies. The central issue to which I shall apply this model are the properties and functions of tones, intervals, and harmonies, although more or less similar models are applicable to other musical aspects as well.[9]

EXAMPLE 1. Relationships of analysis, music theory, and psychoacoustics in analytical explanation



Music analysis is concerned with interpreting the functions of particular occurrences of tones, intervals, and harmonies. Such functions are influenced (but not determined) by theoretical principles, by which I mean here general principles of organization in a certain body of music. For the present considerations, the most pertinent theoretical issue is the normative

[9] In a more general model, the realm of psychoacoustics would have to be supplemented by other studies relevant to the context-independent aspects of auditory perception (including the psychology of memory and *Gestalt* psychology). Some points of comparison for the present model are worth mentioning. The relationship between Bregman's (1990) notions of "primitive" and "schema-based" streaming is roughly comparable with that between the realms of psychoacoustics and theory in the present model; primitive streaming is an automatic perceptual process, whereas schema-based streaming is influenced by attention and learning. Agmon's (1990) model, on the other hand, involves the physical, perceptual, and cognitive domains. Of the realms of the present model, psychoacoustics is concerned with the relationship between the physical and perceptual domains, whereas theory and analysis both belong to the cognitive domain in Agmon's model.

functional status of tones, intervals, and harmonies (e.g., functional consonance and dissonance). Theoretical principles, in turn, are influenced (but not determined) by psychoacoustical aspects, that is, the relationship between acoustical signals such as tones, intervals, and harmonies and their perceptual properties, as uninfluenced by any musical context. Hence, in the present model, psychoacoustical considerations are not directly applied to analysis. Their significance for analysis is indirect, mediated through theoretical principles.[10]

Example 1 also identifies two levels of *contextual* factors. In each case, such factors derive from the treatment of material (tones, intervals, and harmonies) rather than from the properties of the material as evident in the underlying realm in the model (theory for analysis, psychoacoustics for theory). For example, the analytical interpretation of an occurrence of an interval is influenced not only by the theoretical normative status of the interval, but also by the "particular context": the treatment of the interval in the particular occurrence. The theoretical principles concerning the interval, in turn, are not only influenced by the psychoacoustical properties of the interval but also by the "general context": the general treatment of the interval in the relevant body of music.[11] The relative importance of these factors is widely variable according to the musical style and aspect of organization.

Throughout the present studies it is assumed that the criterion for the validity of theory and analysis is their *relevance for or concordance with the musical experience of a sensitive listener*. This means that the extent to which the composers were aware of the theoretical principles and analytical observations is a secondary consideration, which is mostly left out of the discussion. While composers—at least the present ones—count as sensitive listeners of their own music, they are by no means necessarily aware of all experientially relevant aspects in their composition process.[12]

The above criterion is by no means original or unusual. Similar assumptions are held either explicitly or implicitly in most theoretical and analytical enterprises. For example, Lerdahl

---

[10] However, one cannot entirely rule out the direct significance of psychoacoustics for analysis. Psychoacoustical effects inevitably accompany all particular musical events and are potentially significant for their analytical interpretation. However, the present studies focus on prolongational relationships regulated by consistent functions of intervals and harmonies, which makes the theoretical level indispensable. From a general viewpoint, the "thickness" of the theoretical level—the extent to which intervallic and harmonic usages are crystallized to consistent organizational principles—may vary in different musics. In particular, such crystallization may be historically preceded by an experimental phase, in which the future usages are adumbrated less systematically, relying on more immediate psychoacoustical effects. One may see such experimental phases as leading both to "common practice" tonality and out of it. The latter process is evident in the Romantic practices of granting special roles to harmonies such as the major V9/7, which are functionally dissonant in "common practice" but whose perceptual stability is supported by psychoacoustical factors of rootedness, and which adumbrate the kind of post-tonal referential harmonies discussed in **II** (for discussion of this historical process, see section 5.2.3 below, **II**: section 3, and **III**).

[11] The notions of "general" and "particular" context have been designated for the discussion in the present section, written after **I–III**. In the articles and elsewhere in the present text "context" and "contextual" are often used without specifying which type of contextual influence is involved. It is to be hoped that the pertinent meaning will appear from the context (!) of the discussion.

[12] Schoenberg, for example, admits his ignorance of the "laws" guiding his creative process in several occasions, like in Schoenberg 1911/1922/1978: 421: "Laws apparently prevail here. What they are, I do not know."

and Jackendoff (1983: 1) define the goal of music theory as "a *formal description of the musical intuitions of a listener who is experienced in a musical idiom*." (Italics original.) Again, Carl Schachter (1999: 19) observes, "A good analysis is always verifiable by the educated ear." While the choice of words in my own formulation differs from these statements, this does not convey any essential disagreement with either of them. My substitution of "sensitive listener" for Lerdahl and Jackendoff's "experienced listener" and Schachter's "educated ear" is a minor nuance based on my feeling that being "experienced" or "educated" does not always guarantee sensitivity to important musical aspects. On the other hand, even less extensively "experienced" or "educated" listeners may sometimes possess such sensitivity. By "sensitive listener" I wish to leave unspecified the extent to which the faculty of being affected (cognitively and aesthetically) by musical relationships is acquired through sensitization to music by extensive experience and education, as opposed to something like natural talent.

While there is relative agreement that theory and analysis should reflect the "musical experience of a sensitive listener" or "the musical intuitions of an experienced listener," a problematic feature in such criteria is the subjective nature of such experiences and intuitions. It is also not easy to determine who actually qualify as "sensitive" or "experienced" listeners. In the present studies, theoretical and analytical observations are tested against the intuitions of the author, as is customary in theory and analysis. However, Example 1 makes explicit a framework for justifying such intuitions on a more objective or intersubjective basis.[13] All relationships in this model are clearly significant for the experience of sensitive listeners. The ways in which

---

[13] Testing analytical observations against one's intuitions can mean either of the following (or something in between): On the one hand, the analysis may explain cognitive or aesthetic effects—for example, why this or that note sounds "right"—which the listener has sensed without being aware of their explanation. On the other hand, it can actually sensitize the listener to such effects. On the basis of this difference, David Temperley (2001b) has seen a fundamental incompatibility between "descriptive" and "suggestive" theoretical approaches, demanding theorists to state explicitly whether their purpose is "shedding light on our current hearing of the piece" or "enhancing that hearing in some way" (ibid.: 67). I do not fulfill that demand; I leave unspecified which of these goals I wish to attain, since that depends crucially on the listener (if "experienced" and "sensitive" listeners are somewhat ambiguous concepts, Temperley's "our current hearing" is even more unclear). Given that Temperley makes several comparisons between music theory and linguistics, it seems appropriate to illustrate by a linguistic analogy that the distinction between "descriptive" and "suggestive" approaches is less fundamental than Temperley sees. Suppose there is a valid theoretical presentation of Chinese grammar. For the present author, it would be "suggestive," to use Temperley's expression. Studying this presentation would "enhance my hearing," that is, help me learn to understand Chinese. For Chinese-speaking people, the very same presentation would be "descriptive": it would explain features relevant for their current understanding. While there are differences between language and music (an issue is much too complex to be discussed here), it is an undeniable fact that there are different levels of musical understanding (this I can say on the basis of my own listening experiences at different times). Whether theory or analysis is "descriptive" or "suggestive" depends on such a level.
(It is possible that Temperley has emphasized this distinction because he has felt that some branches of theory are exclusively "suggestive"; they concern aspects that can have no effect on hearing unless one expressly takes them into account when listening. Through concern for perceptual and contextual support for theoretical principles, I wish to ensure that the present approach does not fall into that category. It is also possible that Temperley's views reflect the emphasis that his theoretical work, as manifest in Temperley 2001a, lays on "infrastructural" aspects of organization, whose experiential relevance may perhaps be less listener-dependent than that of more "high-level" organization.)

particular occurrences of musical entities (tones, intervals, and harmonies) are experienced clearly depends on both their normative functional status and on their treatment in the particular context. Moreover, the experiential significance of organizational norms concerning such entities depends both on their psychoacoustical properties and on their treatment in the general context.

To justify these statements, and to illustrate the relationships in Example 1, let us first consider the role of octave relationships in conventional tonal organization. The way in which individual octave relationships are experienced is clearly influenced by the general organizational principle of octave equivalence; in other words, theory pertains to analysis. On the other hand, such pertinence may be diminished by the special features of a particular context, which may cause, for example, the octave to function as an embellishment of a seventh, rendering the octave relationship less significant for structural relationships.[14] The theoretical principle, in turn, is influenced by the special psychoacoustic properties of the octave.[15] The choice of the octave as the basis of such a principle is certainly no coincidence. However, the theoretical principle of octave equivalence and the psychoacoustical properties of the octave are by no means the same thing. The organizational significance of octave relationships goes much beyond the direct manifestations of the psychoacoustic special properties of the octave (cf. Agmon 1990: 291–93).[16] A direct application of psychoacoustics to analysis would fail to account for this significance—which illustrates why we need a specific theoretical level to mediate between psychoacoustics and analysis.

One of the arrows in Example 1 remains to be considered in this connection: the influence of "general context" on the theoretical principle. In a sense, we could say that the principle of octave equivalence is brought about by the very "general context" comprising most Western music (and some other musics as well): the treatment of the octave in this "general context" reinforces or "exaggerates" the psychoacoustically supported tendency towards a strong relatedness between octave-related pitches. However, due to the almost unlimited scope of such a "general context," it may be somewhat misleading to speak of "contextual" influence in this case. In the following, I shall employ the notion of *cognitive generalizations and extensions* to identify the reinforcing impact of very general musical practices on the relationship between

---

[14] However, see note 19 regarding such particular contexts.

[15] These special features are based on two main aspects. First, the harmonic spectra of the tones separated by an octave are closely related: the spectrum of the higher tone comprises every other harmonic in the spectrum of the lower tone. Second, the octave relationship itself mimics that between the first two harmonics, the recognition of which is important for virtual-pitch perception (see section 5.2.1).

[16] This becomes evident, for example, by considering multiple octaves. The organizational status of multiple octaves is similar to that of single octaves, even though the former have no special psychoacoustical status comparable to the latter. (Multiple octaves correspond to frequency ratios 2^n. To my awareness, there are no special psychoacoustical properties related to such ratios, as opposed to other integer ratios.)

psychoacoustical and theoretical principles (see section 2.2). By "general contexts" I shall refer to somewhat more clearly demarcated bodies of music.[17]

Such more clearly demarcated general contexts become evident if we consider, instead of octave equivalence *per se*, the more precise forms that this principle takes in the organization of different kinds of music. A central issue in the present studies are the different ways in which octave equivalence is restricted in harmonic organization according to the general context. In the general context of conventional tonality, for example, there is a restriction that bass-related intervals cannot be inverted without altering their functional status, even though they may be enlarged or reduced by octaves.[18] Some of the present examples manifest other kinds of restrictions (section 4.1.2).

For further illustration of the relationships in Example 1, consider the issue of consonance and dissonance, which is a very central issue for prolongation (it is addressed by the first of Straus's conditions). The analytical interpretation of occurrences of intervals and harmonies is influenced by theoretical principles concerning their normative functional consonance–dissonance status in the relevant body of music. As a norm, dissonances are subordinate to consonances but in certain particular contexts occurrences of consonances and dissonances may hold the reverse relationship.[19] Whereas music theory is concerned with *functional* consonance and dissonance, psychoacoustics identifies phenomena underlying *sensory* consonance; two such phenomena: *critical band* and *virtual pitch* are discussed in sections 5.1–2 below. Such phenomena produce a greater or lesser tendency for intervals and harmonies to be heard as stable. Functional consonance is influenced by such tendencies but also by the general context. The less the norms of consonance and dissonance are supported by psychoacoustics, the stronger the contextual means have to be in order to make the norms experientially relevant.[20] For a full understanding of the issue, it is essential to recognize both

---

[17] To illustrate what this would mean in this connection, consider the hypothetical example of music in which some other interval is elevated by contextual means to the status that the octave normally has. That this example is only hypothetical, that it is not realized (to my awareness) in actual music, suggests that for this principle the influence of psychoacoustics is particularly strong.

[18] Particular contexts may, in turn, override general restrictions in octave equivalence. Hence, while the 6/4 chord is normatively dissonant, it may assume a consonant function if the particular context reinforces its association with the 5/3 chord with which it is "octave equivalent."

[19] Actually, such "particular contexts" depend on aspects such as meter and motives which are, of course, not beyond theoretical description, if one does not confine theoretical considerations to pitch-based stability factors as in the present discussion. In the present text, other factors are considered in section 3.2.3.3. However, while allowing for such additional factors enhances the power of theory, that power does not reach the capability of determining analytical interpretation without case-specific considerations of particular contexts. The additional theoretical discussion goes no further than identifying separate principles concerning the impact of different factors on structural relationships. Such principles are involved in the justification of analytical decisions, but how they should be weighted against each other in each occasion is not determined by theory (at least not the present one). Such principles resemble Lerdahl and Jackendoff's (1983) "preference rules," though the present studies treat them in a less formal and less systematic manner.

[20] As an example of a functional consonance system in which perceptual consonance has been overridden by the general context, one may consider some music by Liszt, in which, according to analyses of Robert P. Morgan (1970) and Howard Cinnamon (1986), the *augmented* triad acquires the role of the consonance to which other

the distinctions and the relationships between the realms of analysis, theory, and psychoacoustics.[21]

Above, I have identified music theory as being concerned with the principles of organization in a given "body of music." This expression may need some clarification. It should be noted that there is no specific lower or upper limit for the size of such "bodies of music." The size varies, according to the music and principle involved, from vast cultural entities, such as "common practice" tonality, to individual pieces or sections thereof. Regarding functional norms relevant to prolongation, such as consonance and dissonance, there is a significant difference between tonal and post-tonal music. In tonal "common practice" music, such norms are common to an extensive repertoire, but in post-tonal music, they (if they exist) may apply to individual pieces or parts thereof—even though the present studies also reveal some more general tendencies (**I**: section 5, **II**: section 3, **III**). This implies a considerable difference between the tasks required by prolongational analysis of tonal and post-tonal music. The analyst of tonal music has relatively little need to consider functional norms from the theoretical viewpoint, whereas the analyst of post-tonal music cannot avoid piece-specific theoretical considerations. This difference stems partly from our weaker knowledge of functional norms in post-tonal music but also from the stronger individualization of organizational principles in post-tonal music.

While there is thus no lower limit to the size of a "body of music" relevant to a theoretical principle, this does not remove the distinction between theoretical principles, such as functional consonance and dissonance, and analytical observations, such as the structural order of particular occurrences of harmonies. The minimal requirement for a theoretical principle is general significance in a well-defined stretch of music. For example, even though there are, in conventional tonality, particular contexts in which occurrences of seventh chords are structurally superior to triads, there is no justification for calling seventh chords locally consonant. Such

---

chords, including positions of minor triads, resolve. The contextual means for effecting this are explicit indeed: consider, for example, the opening of the *Faust* symphony.

21 Much confusion has been caused by the failure to recognize this. Some have inferred that since (psycho)acoustical and music-theoretical consonance are not identical, there is no connection between the two. Others, again, have denied the validity of the music-theoretical principle on the grounds that it does not explain each and every analytical decision. For a recent example, Larson's (1997) considerations on psychoacoustical factors of consonance are clearly mistaken (110–11), and his attempts to reject the consonance–dissonance condition of prolongation (Straus's condition #1) on the basis of examples showing exceptions to this condition (128) are exaggerated. But even Lerdahl's (1997) response to Larson does not reach full clarity. Lerdahl rightfully corrects Larson's considerations on psychoacoustical phenomena, distinguishing between sensory (= psychoacoustical) and musical consonance. However, in the discussion of the latter aspect, Lerdahl would appear to confuse normative functional consonance with the structural functions of particular events. He first states that "[m]usical consonance […] is the product of cognitive acculturation to the way intervals behave in specific musical idioms." (151.) He thus defines "musical consonance" as an organizational principle, recognizing the significance of general context for this principle. However, he goes on to note that "[t]here are cases, such as those noted by Larson, in which prolongational context reverses the usual ranking of sensory consonance and dissonance—that is, in which musical overrides sensory consonance." (Ibid.) Here "musical consonance" refers to structural functions of particular intervals, rather than to an organizational principle: Larson's examples do not demonstrate how normative functional consonance overrides sensory consonance, but how particular contexts override normative consonance.

cases do not involve any general principle favoring the superiority of seventh chords over triads even in a limited stretch of music; the superiority of the particular occurrences is not based on such a principle but on the particular context.[22] By contrast, some of the present examples (Scriabin's *Vers la flamme* [**II**] and Debussy's *Ce qu'a vu le vent d'ouest*: [**III**]) show genuine examples of locally differentiated consonance systems. In these examples, different sections consistently relate structural superiority with different—though related—chord types.

According to Example 1, analytical observations may thus be justified by combined considerations of theory and particular context. Theoretical principles may be justified by combined considerations of psychoacoustics and general context (in addition, theoretical principles may be supported inductively by their productivity in analysis). Such manners of justification do not amount to watertight proof. The model in Example 1 implies what kind of arguments may be made in favor of theory and analysis, but does not specify how such arguments should be weighted against each other. (Moreover, for reasons of space, such arguments are not always made explicit in **I**–**III**.) However, allowing for the relationships in the model provides considerable illumination for the evaluation of competing theoretical or analytical claims. In particular, when a theoretical or analytical claim is feebly or not at all supported by any of the relationships identified in the model, it may be unequivocally demonstrated as invalid.[23]

2.2 COGNITIVE GENERALIZATIONS AND EXTENSIONS

2.2.1 GENERAL CONSIDERATIONS AND EXAMPLES

The notion of cognitive generalizations and extensions has been introduced in connection with octave equivalence. More generally, I refer by these notions to situations in which the organizational status of a musical entity (interval or harmony) holds a correspondence with the psychoacoustical effects induced by the entity in some simple, typical, or "basic" circumstances but prevails also in circumstances in which the psychoacoustical effects are different.

Some kind of cognitive generalizations are an inevitable part of any consistent interval functions, since the psychoacoustical properties of intervals greatly vary according to

[22] Or, more precisely, on the influence of other "parameters," as discussed in note 19.

[23] **I** offers some such demonstrations. Concerning theory, the inadequacy of the set-theoretical approach to harmony and intervals (i.e., that neglects effects of registration) for issues of prolongation is demonstrated by considerations of both psychoacoustics and the general context: registral distinctions—such as those between complementary intervals—crucially affect both the psychoacoustical properties of harmonies and intervals and their general treatment in the music by Schoenberg considered in **I**. Concerning analysis, some of the voice-leading relationships in Roy Travis's (1966) analysis of Schoenberg's op. 19/2 fail—notwithstanding the value of his analysis as a pioneering attempt—to be supported by either a consistent theoretical background or by the particular context (**I**: 251).

circumstances—such as register, spectrum, and sound/noise level—that have little or no impact on their organizational functions.[24] Octave equivalence brings about further cognitive extensions. Owing to the principle of octave equivalence, the music-theoretical conception of harmonies and intervals is somewhat different from (psycho)acoustical. In conventional tonality, for example, bass-related intervals retain their theoretical functional status regardless of octave enlargements and reductions. Whereas the functional dissonance of, say, the major seventh has a correspondence in sensory dissonance (roughness), enlarging the interval by several octaves removes the roughness but not the functional dissonance.

Another cognitive extension is produced by associations between temporal successions of tones and simultaneities. The psychoacoustical phenomena pertinent to successive and simultaneous tones are quite different, but in musical organization arpeggiated harmonies share the functional status of block chords.

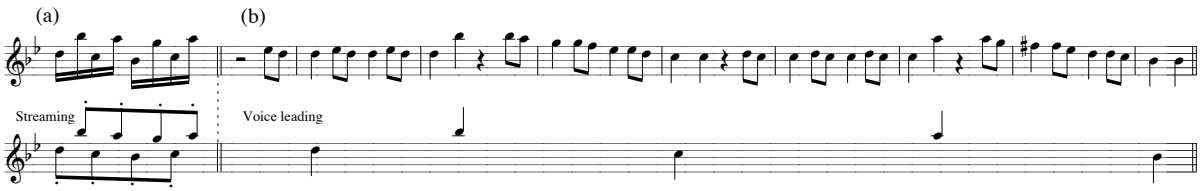
The dimension of temporality involves a further cognitive extension, based on the association between slower and quicker tone successions. A basic principle of voice leading, manifest both in conventional tonality and in the present examples, is the *proximity principle of voice leading*, under which small melodic (=horizontal) intervals, usually semitones and whole-tones, function as voice-leading intervals (melodic connectives with no arpeggiating function) and larger ones function as arpeggiations (section 3.2.3.4). This principle gains psychoacoustical support from the significance of pitch distance for *auditory streaming*(section 5.3). Consider the passage in Example 2a. Successive tones are separated by large intervals, but the tones belong alternately to two stepwise lines. If the tempo of presentation is rapid, these lines are automatically segregated to two auditory streams. Like stream segregation, the proximity principle of voice leading is based on pitch distance, but it involves considerable temporal extensions. For example, the melody in Example 2b implies two voices, even though the tempo in which the notes within these voices follow each other is much too slow to evoke automatic stream segregation in the psychoacoustical sense. The segregation of the melody in Example 2b into two voices is thus cognitive rather than psychoacoustical, but it is supported by the association with passages such as Example 2a that induce automatic stream segregation.[25]

---

[24] The rules of traditional voice leading, such as the prohibition of parallel fifths, may also largely be understood as brought about by cognitive generalizations and extensions of psychoacoustical phenomena. For a recent discussion, see Huron 2001.

[25] For a general presentation on auditory stream segregation and a discussion of its musical implications, see Bregman 1990.

EXAMPLE 2. Auditory streaming and voice leading.



All these examples of cognitive extensions—octave equivalence, arpeggiation, and temporal extensions of voice-leading relationships—may be described as being based on our ability to associate somehow "extended" versions of intervals and harmonies with those in more "basic" circumstances. Such cognitive extensions presuppose learning: presumably they are developed by listening to music to which they pertain (cf. the observations on general context above). However, the learning process is crucially assisted by the psychoacoustical properties of the "basic" cases. It is certainly easier to learn to hear music in terms of octave equivalence than in terms of equivalence based on any other interval. And while the functional significance of small intervals (seconds) in voice leading has to be learned for all except the most minute temporal spans, such a principle is much easier to learn than one based on some other intervals (say, tritones).

A more detailed description of such learning processes and their ramifications in music history and individual psychology is beyond the present concerns. Suffice it to note that octave equivalence, arpeggiation, and temporally extended voice-leading relationships (as in Example 2b) are well-established basic phenomena in musical organization. Their cognitive reality, or "relevance for the musical experience," is not in doubt. However, setting the limits for the range of their applicability poses more difficult problems. As already observed, one of the central concerns in the present studies is the need to put restrictions on octave equivalence in the interpretation of harmonies and intervals (section 4.1). Another crucially important question is whether some kind of limit should be posited to the temporal extensions of voice-leading relationships. In both conventional Schenkerian theory and in the present studies, there is no such limit. These approaches assume that similar principles of voice-leading, based on arpeggiation and voice-leading intervals, apply at different hierarchical levels of structure, involving unlimited temporal spans. This is one of the aspects of Schenkerianism that have led its critics to doubt the experiential relevance of its notions, since our capacity to grasp voice-leading relationships seems to strongly depend on time scale. This difficult and central issue cannot be exhaustively discussed here; the focus of the present studies is on the generalization of Schenkerian principles to works outside conventional tonality, rather than on the justification of Schenkerianism. However, the above definition of theoretical and analytical validity with its concern for experiential significance warrants some considerations of the issue. I shall present

such considerations in the next subsection, illustrating them by two musical examples, one tonal and one from the present post-tonal repertoire.

Prior to these considerations, one interesting phenomenon relevant to cognitive extensions should be pointed out. The experiential significance of cognitive extensions does not always rely solely on the listener's general knowledge of such extentions; the association between "extended" and "basic" circumstances is often "concretized" by musical features that particularly refer to such association. For an example of "concretization" of octave equivalence, one may consider the treatment of D in Debussy's *Voiles*; see Example 4 below.[26] As discussed in **II**, the structural significance of D relates with its status as a *root support* of the bass B♭, a property stemming from its correspondence with the fifth harmonic; see sections 4.2 and 5.2 below (Example 4a shows chord "U" comprising approximations of harmonics 1, 5, 7, 9, and 11 of B♭1;[27] the numbers in the Example show bass-related intervals in semitones, modulo 12). The main structural occurrence of D is in the high register, octave 7, which clarifies D's structural importance but takes it outside the range of harmonics directly relevant to rootedness. The root-supporting status of D7 relies thus on its octave relationship with D4, the due approximation of the fifth harmonic of B♭1. This relationship is concretized by two significant features: the parallelism between the figures that introduce a salient (but non-structural) foreground D4 and the eventual D7 (mm. 9–10 and 15–22; see Example 4e) and the salient return of the D4 immediately after the D7 (m. 22; Example 4c).

The present studies also include several examples in which arpeggiations are followed by block-chord presentations of similar harmonies, concretizing the connection between the two. One example is discussed in reference to Example 7c below.

Finally, both conventional Schenkerian analyses and the present analyses frequently reveal situations in which a large-scale voice-leading relationship is concretized by association with a similar relationship that is closer to the surface and thus closer to the time scale relevant to auditory streaming. Such interlevel relationships take different forms, involving different aesthetic effects. For example, a small-scale progression may occur during the opening element of a large-scale progression, "preparing the ear" for the latter. Or the small-scale progression may occur towards the end of the large-scale event, "confirming" it or "reminding" of the starting-point of the progression. The small-scale and large-scale events may also occur in temporally separate spans, producing "concealed repetitions."[28] All these informal expressions describe some kind of supporting effect that the small-scale correspondence lends to the experiential significance of the large-scale event (for a sensitive listener).

---

[26] Example 4 reproduces, with slight adaptations, parts of Examples 15, 17, and 18 of **II**.

[27] The present text numbers octaves according to the notational standard of the Acoustical Society of America. Octave 4 ranges from the middle C to the B above it, corresponding to the conventional one-line octave; other numbers refer accordingly to octaves below and above. This notation is also used in **I** and **II**, whereas **III** uses traditional notation (in which c^1 is middle C).

[28] Concealed (or hidden) repetitions are widely discussed in Schenkerian studies; see, for example, Burkhart 1978.

2.2.2 VOICE LEADING ON DIFFERENT TIME SCALES

The idea that voice-leading relationships apply hierarchically at different levels of structure lies at the core of Schenkerian theory and prolongational theory in general. Without any supporting factors, relationships between temporally distant tones are unlikely to be experientially relevant. Prolongation plays a crucial role in supporting such relationships. The basic idea of prolongation is that harmonies and tones may be experienced as governing extended spans of music while not being literally present all the time. Thanks to this phenomenon, two temporally distant tones may govern consecutive spans. The distinctive features of prolongation are demarcated by the pitch-based conditions that are discussed in section 3.2, on the basis of Straus 1987. Apart from these conditions, prolongational structures are supported in widely variable ways by correspondences with additional factors concerning either temporal articulation or associations between tones.[29] Such factors include changes of figuration and texture, formal (thematic) articulation, rhythmic, registral, motivic and gestural relationships, and relationships between different scales of organization.

Despite such supporting factors and the productivity of Schenkerian analysis, the experiential significance of large-scale voice leading has often been questioned. Enlarging the temporal scale has been seen as rendering voice-leading relationships increasingly difficult and finally impossible to perceive. The study which is perhaps most often cited as supporting such a view is Cook 1987. Cook conducted listening experiments to examine whether and how large-scale tonic returns affect the listeners' musical experience. Two versions of tonal pieces were presented to a sample of listeners consisting of "first or second year music majors at the University of Hong Kong" (ibid.: 200). One of the versions was the original version of a piece of classical music, which returns to the original key at the end, in accordance with the normal practice of tonality. The other version was modified so as to end in a different key. The subjects were questioned which version better manifested aesthetic qualities such as coherence, completion, pleasure, and expressiveness. The results showed no significant preference for the original version with respect to any of these qualities, except for the very shortest examples. According to Cook, these experiments "suggest that the influence of tonal closure over listeners' responses is restricted to a maximum time scale, possibly on the order of 1 min." (Ibid.: 203.) Since prolongational analysis often involves much longer time scales and since the character of voice-leading relationships is often more subtle than simple tonic returns, such a conclusion casts doubt on the experiential significance of large-scale voice-leading relationships in general.

---

[29] While Straus (1987) rightfully emphasizes the distinction between prolongation and association, these aspects are thoroughly interrelated. (For a demonstrative example, see the analysis of Webern's op. 3/1 in **II**.)

In assessing the significance of Cook's study, it should first be noted that the circumstances in these experiments were hardly conducive to ideal listening sensitivity. Even if we leave aside the question whether the status as a first or second year music major is a guarantee of well-developed listening abilities, there are problems caused by the experimental arrangements and, above all, by the subjects' utterly insufficient exposure to the music: each version was played just once (in two experiments, separated by one year; the order of the original and the modified version was reversed in the second experiment).[30] All in all, these experiments are far from adequate to justify the conclusion that tonal large-scale relationships are irrelevant to the musical experience, if we are primarily concerned with somewhat deeper experiences, as has traditionally been assumed in analytical studies.

Despite these reservations, Cook's study is not devoid of important implications for theory and analysis. While the experiential significance of tonal large-scale relationships may be greater in more favorable circumstances, Cook's experiments illustrate that such relationships tend to play a less obvious and less salient role in the perception of large-scale organization. Even if the role of such relationships is strengthened in more sensitive or "deeper" listening, they supplement rather than displace those factors that presumably dominated the large-scale experience of Cook's subjects. If nothing else, we may thus take Cook's study as a healthy reminder that experientially significant relationships of large-scale voice leading do not arise solely from the presence of the tones involved by these relationships but crucially depend on supporting factors, such as those enumerated at the beginning of this subsection. These considerations may be compared with Cook's own. While he does not seem to share the above-discussed reservations about the quality of listening experience in these experiments, he notes that their implications are less damaging for the theoretical notions of large-scale pitch relationships than one might think:

> […] although it [the tonal plan of a sonata] may not be directly perceived by the listener, its effects will be everywhere apparent in the music. The disposition of textures and thematic materials, the patterning of loud and soft passages and of high and low tessitura, the pacing of tension and relaxation—all these aspects of a sonata are organized around the tonal plan and serve to project its structural closure in a directly perceptible manner. Hence if large-scale tonal relationships are not in themselves audible, that does not necessarily mean that they are of no musical significance: it may just be that their influence on what is heard is an indirect one. It may be helpful to draw an analogy with a conductor's gestures, which are indispensable in coordinating the performance of a complex orchestral work, but which do not need to be visible to the listener in order for him to respond aesthetically to the music. (Ibid.: 204–5.)

Cook refers here to factors similar to those listed above as having the function of supporting large-scale relationships. But whereas I observed that large-scale pitch relationships are highly dependent on such factors, Cook says, in effect, that the latter totally outweigh the former with respect to experiential significance. With respect to their practical consequences,

[30] For another kind of criticism of Cook's methodology, see Gjerdingen 1999.

these views are actually not so far from each other. Under both views, prolongational large-scale structures may relate meaningfully with the musical experience. According to the present view, such a relationship derives partly from the notes involved and partly from a host of supporting factors. According to Cook, it would seem, the relationship derives solely from these supporting factors; the notes in large-scale structures function only as metaphorical representations of other musical aspects.[31] This difference may also be expressed in terms of the above criterion of analytical and theoretical validity (section 2.1). Under the present view, voice-leading structures are regarded—at least partially—as "relevant to" the musical experience, whereas Cook's view would be more compatible with the weaker expression "concordant with." In both cases, large-scale voice-leading analysis is able to meet this criterion, but, of course, the motivation for employing such an approach is enhanced if one accepts the present view.

To get closer still to a reconciliation between these views, it may be observed that they actually represent two points in a continuum of possibilities concerning the relative significance of pitch relationships as such vis-à-vis supporting factors for the experience of voice-leading relationships. Cook's view of large-scale structure represents one extreme in this continuum, whereas the other extreme would be a situation in which pitch relationships are as such sufficient for generating structural relationships. The present view of large-scale structure is somewhere in between; where exactly it lies varies from case to case. Sometimes piece-specific features highlight background pitch relationships in a more or less directly perceptible manner, sometimes such features are less evident, in which case we get closer to Cook's position. A similar continuum also pertains to listening experiences of different people at different times. Moreover, it pertains to different scales of organization. At the surface level, tone relationships need the least amount of supporting factors to generate experientially relevant voice-leading relationships, whereas proceeding to "higher" levels increases the significance of such factors, even if not quite nullifying the significance of pitch relationships.

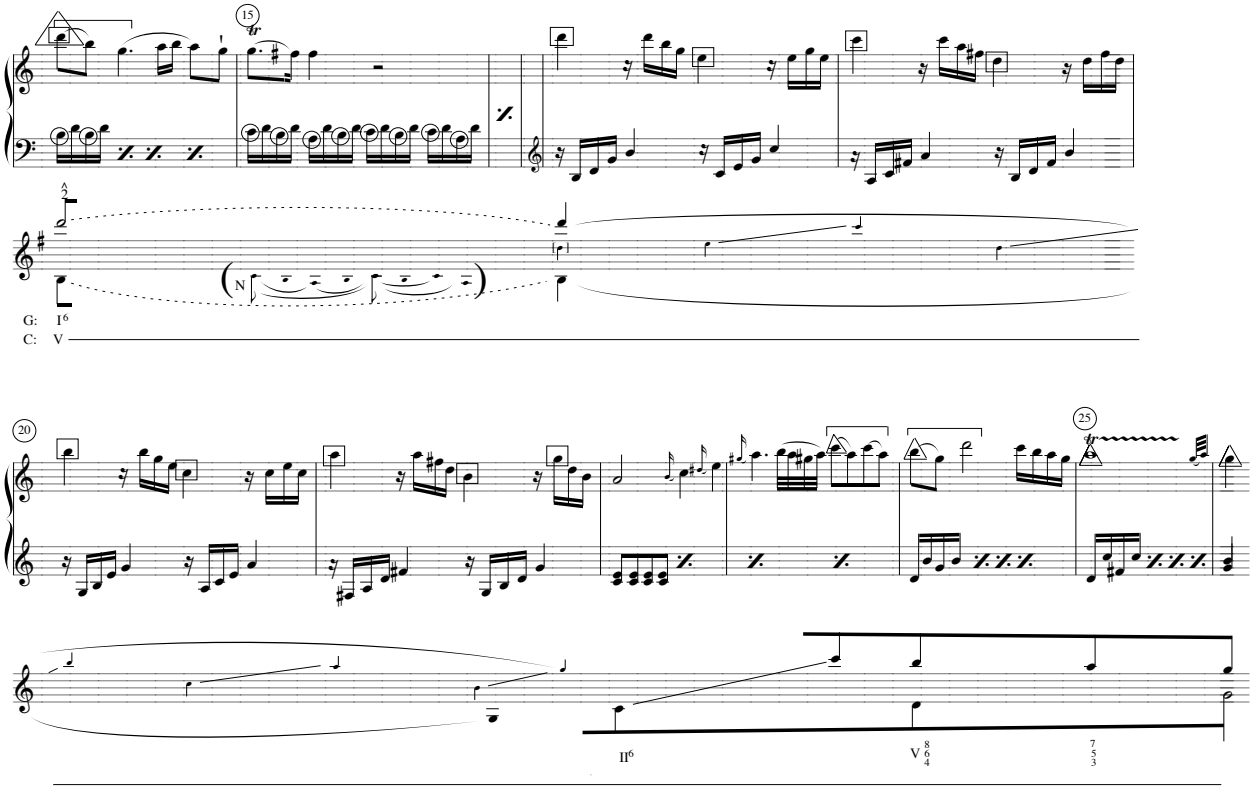
The preceding discussion has made clear that it is difficult to describe the experiential significance of large-scale voice leading from a general viewpoint since such significance crucially depends on supporting factors that vary from case to case. Thus it seems appropriate to present a couple of case studies to illuminate the issue. In the following, I discuss the subsidiary theme of Mozart's Sonata K. 545, first movement (Example 3),[32] and Debussy's *Voiles* (Example 4). In the Mozart example I focus on the way in which the factors relevant to voice-leading relationships and the concomitant experiential effects are transformed by enlarging the structural and temporal scale. In the Debussy example, I concentrate on one large-scale connection and factors supporting it, offering a foretaste of the analytical observations in the present studies. Together, the two examples also provide material for a comparison of

[31] Cook 1989 (125 ff.) actually discusses the metaphorical implications of voice-leading analysis.

[32] This movement is analyzed in Schenker 1935/1979 in Figures 47,1 (overall structure); 88c (subsidiary theme, i.e., the present passage); and 124,5a (opening). The present discussion agrees with Schenker's readings. For discussion of alternative interpretations, see Snyder 1991.

factors that support structural relationships in conventional tonal music and in the post-tonal repertoire discussed in **I–III**.

EXAMPLE 3. Mozart, Sonata K. 545, first movement, subsidiary theme



In the Mozart example, four structural levels may be distinguished. In the following, I number these levels from the foreground to the background, in contrast to the customary Schenkerian practice. I do not describe the first two levels comprehensively, but only consider some demonstrative samples.

The first level involves the very surface and corresponds closely with auditory streaming. The most demonstrative example of such correspondence is the left-hand figuration in mm. 14–15. This figuration implies two voices, a stationary D4 and a lower voice indicated by circles in Example 3. The segregation between these voices concords more less closely with auditory stream segregation.[33] (For the structural relationships within this voice, see the small notes in the analytical graph on the lowest stave of Example 3.)

[33] The streaming effect may be somewhat ambiguous due to the registral proximity of the two voices; there is even a "stepwise" interval C4–D4 between the two voices. On the other hand, the performer will probably take care of bringing the circled notes out by dynamic emphasis.

At the second level, a bit further from the surface, voice-leading corresponds with patterns that are not auditory streams but that may be characterized as cognitively obvious. Such patterns are exemplified by the two voice-leading strands formed by the rectangled notes in mm. 18–21. The upper strand (D6–C6–B5–A5–G5) continues a voice that remains stationary at D6 at the beginning of the theme (mm. 14–17). Whereas in the preceding example the two voices (mm. 14–15, left hand) were induced purely by the configuration of tones, the present example involves some supporting factors, such as the metric and rhythmic emphasis and local consonance.

At the third level, voice leading becomes less obvious for cognition and more analytical labor is required to reveal it. Nevertheless, it sheds light on the coherence of the theme in a way that is not too difficult to grasp for a sensitive listener. This level is represented by the triangled upper-voice line D6–C6–B5–A5–G5 that spans through the theme. In comparison with the preceding examples, the analytical justification of this level of voice leading involves a much greater set of supporting factors:

> *Clarity of prolongation*. As illustrated by the analytical graph (lowest stave), the upper-voice progression is supported by a harmonic structure that fulfills the normative requirements of conventional tonality.
> *Textural and rhythmic changes*. The texture and rhythm of the accompaniment in mm. 22–23 differs from its surroundings, signaling the role of the II^6 in triggering the structural motion.
> *Registral positions*. D6 (m. 14) and C6 are the highest tones within the spans governed by the I^6 (mm. 14–21) and II^6 harmonies (mm. 22–23). (A similar explanation does not apply to the ensuing B5 since m. 24 includes a higher tone D6, but other factors compensate for this: the straightforward correspondence between metric and structural weight in mm. 24–26, the parallelism between the eighth-note figures C6–A5 and B5–G5 in mm. 23–24, and, of course, the organizational norm favoring stepwise voice leading.)
> *Temporal positions*. All tones in the upper-voice progression occur at either end of the span governed by the supporting harmony. D6 occurs at the beginning of I^6; C6 occurs at the end of II^6; the remaining B5, A5, and G5 occur regularly at the beginning of the accompanying harmonies.
> *Motivic relationships*. The D6, C6, and B5 are introduced by motivically related figures, as bracketed above the musical example.
> *Interlevel relationships*. The overall top-voice progression D6–C6–B5–A5–G5 is reproduced in mm. 18–21 (the upper strand of rectangled notes) and, on the most minute scale, at the end of m. 24. The former reproduction may be depicted as a preparation for the main progression, the latter as a reminder of its starting point (D6) just prior to its completion.[34]

The fourth and ultimate level of organization involves the status of the top-voice D and the governing G major harmony in the overall background structure. They represent the $\hat{2}$ and V in the *Ursatz*.[35] At this level, the organizational significance of voice-leading relationships becomes more difficult to grasp (some might say that we are proceeding from the perceptual to

[34] Since such fifth-progressions occur so frequently in tonal music in general, one may question the specific significance of these interlevel relationships. Observe, however, how the latter is enhanced by its link with the bracketed motivic relationship.

[35] According to Schenker (1935/1979: Fig. 41,1; 124,5a), the *Urlinie* begins from $\hat{3}$. However, even if we were to prefer a reading that starts from $\hat{5}$ (m. 1) and descends shortly to $\hat{4}$ and $\hat{3}$, this would not essentially affect the present viewpoints.

the conceptual). To be sure, the structural position of the $\hat{2}$ is supported by means partly similar to those discussed above. Its appearance is highlighted by the tonicization of (= modulation to) the dominant, by an introduction of a new theme, by its registrally extreme position (D6), and by the emphasis given by its repeated occurrences at metrically strong points in mm. 14–18. Due to the combination of all these factors, it is clearly "concordant with" the musical experience to identify the D6 as a major structural event with a passing character in the overall organization; hence the Schenkerian reading the D6 as the $\hat{2}$ in the *Urlinie* has at least metaphorical validity. Whether the actual tone relationships between $\hat{3}$ and $\hat{2}$ and between $\hat{2}$ and $\hat{1}$ are also "relevant to" the experience of the large-scale organization is a somewhat more difficult question. I shall only consider the first of these relationships. Features that may seem to detract from its experiential relevance include the registral non-correspondence of the $\hat{3}$ (E5) and the $\hat{2}$ (D6) and the lack of emphasis on the former in comparison with the latter. However, the linear connection between $\hat{3}$ and $\hat{2}$ is reflected by interlevel relationships, by a subtle "concretization" of the background connections at the foreground. The $\hat{3}$ is introduced by a A5–G5–F5–E5 line in mm. 3–4. As observed by Schenker (1935/1979: Fig. 124,5a), this is followed by the repetition of this line (with different figuration) and an emphatic continuation to D5 (mm. 5–9).[36] The subsidiary theme retains D at the top voice, but the new harmonic context, thematic entry, and register underline the "elevation" of D to a higher structural level. There is thus a chain of events which first clarifies the $\hat{3}$–$\hat{2}$ relationship at the foreground and then the relationship between $\hat{2}$s on different levels. This chain of events involves rather subtle means, and becoming conscious of it requires some analytical labor. Nevertheless it is by no means beyond "relevance for the experience of the sensitive listener" at an unconscious level.

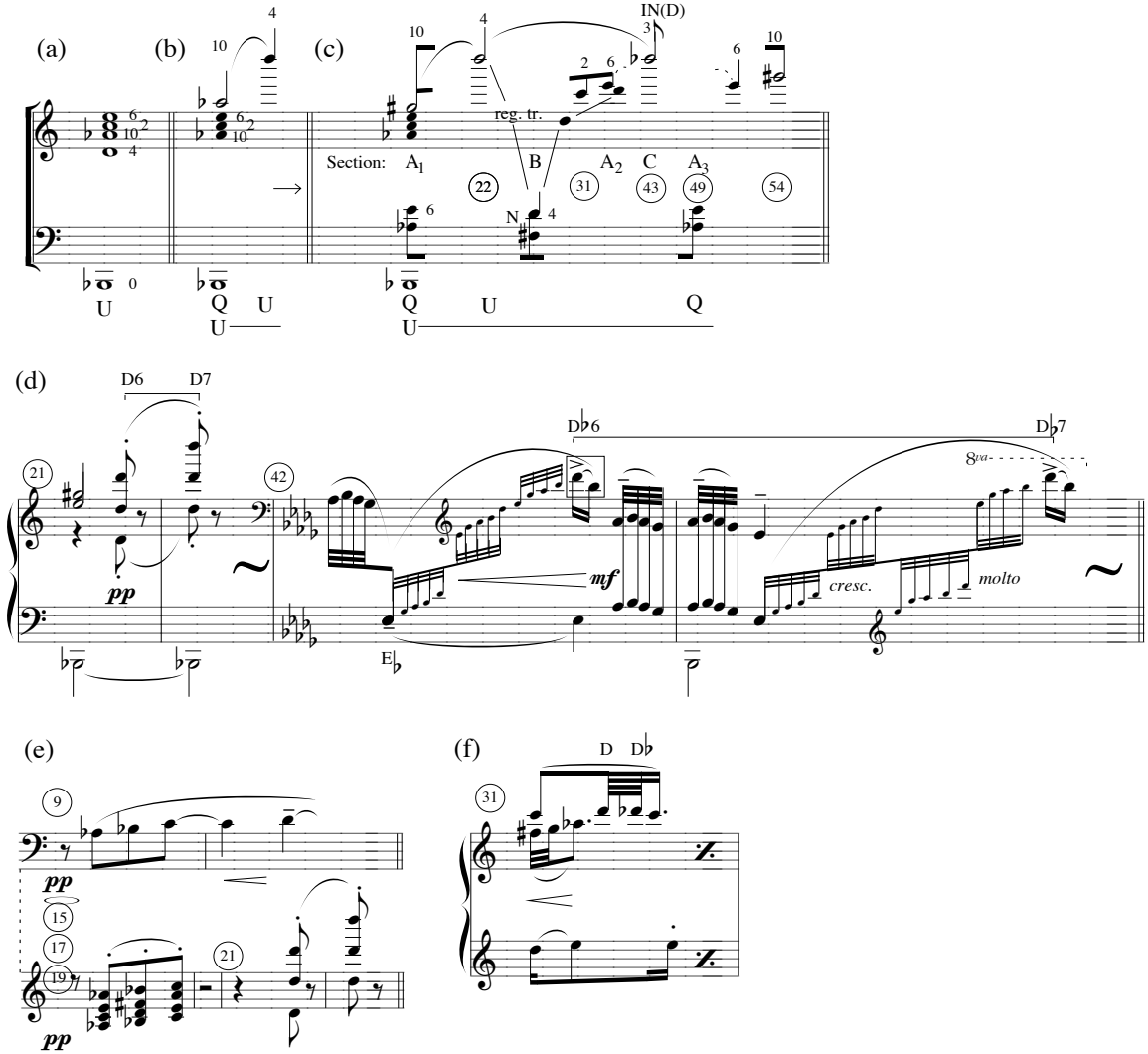
The above observations of the Mozart example illustrate features of voice leading on different scales in a rather typical manner. Near the surface level successions of tones require little supporting factors for inducing experientially relevant voice-leading relationships; further from the surface, such supporting factors become increasingly significant. At the very surface, voice-leading relationships are perceived in a half-automatic fashion, whereas further from the surface, more cognitive activity is involved. At the background level, even the cognitive significance of voice-leading relationships may be less clear, but the voice-leading model coincides with experiential effects at least in a metaphorical sense; moreover, the significance of background voice leading is borne out by interlevel relationships.

I shall now turn to Debussy's *Voiles*, concentrating on the large-scale voice-leading connection from D7 (m. 22) to D♭7 (m. 43). (Foreground events in this prelude are discussed in reference to Examples 15–16 in **II**.) Example 4a–c reproduces structural graphs from Example 17 of **II** in a slightly adapted form. The Example contains some unfamiliar symbols. "IN(D)" indicates a "dimming tone," an element of embellishment identified in **II**; the dimming tone

[36] For an illuminating description of these events, see also Aldwell and Schachter 1989: 4–5.

may be defined as an incomplete neighbor that follows the main tone and lies a semitone lower. Numbers indicate bass-related registrally ordered intervals by semitone notation; Q and U are chords consisting of such intervals, as evident from Example 4a–b.

EXAMPLE 4. Debussy, *Voiles*. Illustration of musical relationships



The D7–D♭7 relationship is supported by a host of factors largely comparable to those of the Mozart example:

> *Clarity of prolongation.* For considerations of harmonic stability, the reader is referred to the analysis in **II**. The means by which D7 is prolonged are, however, straightforward enough to be understood without such considerations. They consist of register transfers: D features constantly in lower-register ostinato figures until the appearance of D♭7 (Example 4c).
> *Coordination with formal articulation.* In the clearly articulated ABACA scheme (Example 4c), D7 occurs in the cadential octave gesture that ends the first A section; D♭7 is the apex tone in the quasi-glissando gestures that open section C. These gestures are shown in Example 4d.

> *Changes of pitch collections*. The structural decisiveness of the D♭7 is underlined by a change from the whole-tone set to the pentatonic set. (The effect of this factor is roughly comparable to that of modulation in tonal music.)
> *Registral position*. D7 and D♭7 are highlighted by registral isolation. They are the two highest pitches in the prelude, and no pitch close to them occurs in the temporal span between them.
> *Gestural parallelism*. As indicated in Example 4d, the outward contrast between the gestures introducing the D7 and D♭7 links with a remarkable parallelism between the octave shifts D6–D7 and D♭6–D♭7.
> *Interlevel relationships*. Two small-scale events are especially relevant. The first of them, the connection between the ways of introducing D on the small and large scale (Example 4e), was already discussed with respect to the concomitant octave relationship. With respect to structural and temporal scales, the small-scale D "prepares ground" for or functions as the "seed" of the large-scale D—to use informal but apt organicist expressions. Second, there is also a surface correspondence for the D♮–D♭ relationship, in m. 31 (Example 4f). The significance of this minute detail is enhanced by its being the single occurrence of foreground semitones.[37]

Taken together, all these considerations show how integrally the large-scale D♮–D♭ relationship ties in with various features of compositional design. One might say that such features appear as if they had been created to clarify the large-scale connection. Findings such as the present ones point to what is perhaps the primary motivation behind large-scale voice-leading analysis. While the experiential significance of large-scale connections is difficult to describe on an *a priori* basis, case studies of "masterworks" show that, in the music of "great" composers (such as Mozart and Debussy), such relationships often tie in integrally with various more readily perceptible features, supporting *a posteriori* their significance.

This does not mean, of course, that owing to all the supporting factors, the "sensitive listener" experiences the large-scale relationship in a same way as small-scale voice leading. There remains a phenomenal difference between different time scales. What has motivated and fascinated prolongational analysts is the evidence that, *despite* this phenomenal difference, there is a tendency in certain art music towards a profound relatedness between organization on different scales, towards bridging the difference. Whereas conventional Schenkerian analysis has gathered evidence for such relatedness in the "masterworks" of conventional tonality, the present studies attempt to do the same for some post-tonal music. As regards the experiential significance of a large-scale connection such as the D–D♭ motion in *Voiles*, it should be observed that not even "sensitive listeners" necessarily "perceive" that relationship in a conscious manner (whereas, at the other end of structural and temporal hierarchy, the perception of the same relationship in the sixty-fourth-note figure in m. 31 derives from automatic stream integration without calling for any musical "sensitivity"). Nevertheless, the "sensitive listener" will experience the all-important aesthetic consequence of the large-scale relationship: that the

---

[37] In contrast to the Mozart example (note 34), it can be easily shown by probabilistic considerations that this interlevel connection is unlikely to be a chance product. On both the small and large scale, there is a single occurrence of semitones. In both cases, the most salient (top-voice) semitone motion involves the same pitch classes (D♮ and D♭), of which the probability is 1/12.

appearance of the D♭ and the concomitant shift to pentatonicism in m. 42 are not arbitrary whims but connect "organically" and "coherently" with the rest of the piece.[38]

As discussed above, a problematic aspect in the above criterion of theoretical and analytical validity involves the identification of "sensitive listeners." Cook's arguments are weakened by the lack of evidence that his subjects' listening experiences were even close to ideal sensitivity. If one attempted to verify the above-discussed aesthetic effect in *Voiles* by listening experiments, one would encounter the same problem (among many others). Instead of trying to solve problems concerning listening experiments, I would raise the question whether the analytical findings do not themselves speak for the experiential relevance of the large-scale connection. After all, there is one prime candidate for a sensitive listener—at least in Debussy's case—namely the composer himself. And it is difficult to explain how all the features supporting the D–D♭ relationship could have come into being, unless they were significant for Debussy's musical experience.

Of course, one cannot generally argue that all musical features are relevant to experience just because the composer has incorporated them into his/her pieces. Large-scale tonic returns are a case in point: it is quite plausible to argue that some or all tonal composers employ such returns, not because they are sensitive to any concomitant experiential effects, but simply because they have chosen to follow the general practice. However, the D♮–D♭ relationship in *Voiles* differs from large-scale tonic returns in two important respects. First, despite all the clarifying factors, the former relationship is much more subtle: analytical considerations are required to reveal it, whereas tonic returns are usually obvious for anyone writing or reading a score (if not for all listeners, as Cook's experiments show). Second, the D♮–D♭ connection is a unique compositional feature whereas tonic returns are a standard device. Debussy was not following a general practice but creating unique relationships. It is possible, and perhaps probable, that he was not consciously aware of these relationships but was responding instinctively to the compositional possibilities that occurred to him. But whatever the level of awareness, it seems difficult to see what could have motivated him to create such an intricate set of relationships for unique use, if not the resultant aesthetically satisfying experiential effects.

While the above discussion certainly leaves much to be clarified concerning the highly complex issue of large-scale voice leading and its experiential effects, it offers the grounds for some important conclusions. Voice leading may largely be understood as being based on the cognitive generalization and extension of auditory stream perception. Moderate temporal extensions will be grasped relatively easily, but greater temporal distances tend to cause greater

[38] The illumination of the "organicism" and "coherence" in the "masterworks" of "great composers" represents a very traditional kind of analytical enterprise, which has been subject to heavy criticism in much of recent musicological discussion. For whatever merits such criticism may have in pointing to the limitations of such an enterprise, it should not make us to lose sight of the fact that qualities such as "organicism" or "coherence" are, from the experiential standpoint, very real qualities. For those who feel that such qualities have already been overemphasized in analytical literature, my invocation of them may perhaps be excused on the grounds that Debussy is not part of the Austro-German tradition primarily associated with them.

difficulties for direct perception. At least partially, such inherent difficulties may be compensated for by prolongational clarity and additional associational factors. Nevertheless, it is possible that towards the background the relevance of voice leading tends to diminish in relation to other factors. However, even insofar as this is true, voice-leading analysis is still able to produce results "concordant with" the experience, if less "relevant to" it. The extent of such relevance cannot be determined in a general manner because the effects of large-scale voice leading depend on case-specific factors. However, at least sometimes such relevance can be strong indeed, as suggested by analytical findings of large-scale relationships that have important ramifications for all aspects of the music, including immediately perceptible ones.[39] Such findings support the view that "great" composers (Mozart and Debussy, for example) were sensitive to large-scale pitch relationships as aesthetic resources and exploited them to a greater or lesser extent—enough to oblige us to be alert for such relationships. Therefore it is prudent not to impose *a priori* temporal limitations to voice-leading analysis, but to push the analysis as far as one is meaningfully able to do, and assess, *a posteriori*, its significance for individual cases.[40]

2.3 THE SPIRIT OF THE PRESENT APPROACH

In most respects the aims and methods of the present studies may be described as characteristic of customary theory and analysis. Its main concerns are to make close analytical readings of musical organization in individual works and to discuss and develop theoretical principles relevant to such analysis. The analyzed repertoire consists of music by acknowledged masters. The traditional, some might say "old-fashioned," notions of "organicism" and "coherence" strongly pertain to the analytical attitude.

An aspect that deviates somewhat from customary theory and analysis is the explicit attention on psychoacoustical, or in broader terms, perceptual aspects. This does not mean that my point of view comes close to that of perceptual studies. Perceptual studies have a role to play only as background factors of basic theoretical principles. Moreover, the only studies I invoke concern psychoacoustics rather than music perception; they are studies of the very basic level of

---

[39] These considerations correspond, I think, with views more or less widely held among analysts. Consider, for example, the following two statements by Carl Schachter: "Of course the deeper levels of structure, by definition, are not as readily accessible to direct perception as are events of the foreground […]" (Schachter 1999: 19). "In tonal music the underlying pattern is not always hidden from immediate perception, but frequently 'is a direct part of musical experience' [a quotation of Charles Rosen]. And this is true of background relationships as well as of the middleground one that Rosen cites." (Ibid.: 35.) Hence (according to the first statement), background tends to be less directly perceived, but (according to the second) there are "frequent" exceptions to this. For a thoughtful discussion of the relevance of organizational principles on different scales, see also Larson 1997: 114–7.

[40] The analyses in **I–III** do not make such *a posteriori* assessment explicit, partly for reasons of space, and partly due to the difficulty of the subject.

auditory perception outside musical context. Nor does my methodology approach that of cognitive studies, even though I am concerned with making theoretical principles as explicit as possible (see, for example, section 3.2.3.3 below), and I do believe that the present principles have a significant correspondence with cognitive structures (of sensitive listeners). In principle, these principles might be formalized along the lines of Lerdahl and Jackendoff's (1983) cognitive approach, but I do not take such a step here.

While perceptual studies thus have a role to play at the starting point of the chain of reasoning leading to analysis, as modeled in Example 1, I have reserved myself the customary theorist's and analyst's right to make the rest of the steps in this chain without explicit considerations on perceptual verification for each of them. The desire to do this is motivated by the desire to proceed to musical aspects more complex and sophisticated than what is usually possible in perceptual studies that are burdened by such verification.[41] On the other hand, by considering the relationships between basic theoretical concepts and auditory perception, I have aimed to avoid the danger of basing theory and analysis purely on complex and sophisticated speculation, with no concern for a meaningful relationship with perception (That this danger is real is unfortunately evident in much recent theoretical literature.)

Despite the interest in psychoacoustics, the share of psychoacoustical considerations in the argumentation of the present studies is rather limited. There are, in all, four theoretical principles justified partly on psychoacoustical grounds. The first is the significance of registration for the identification of harmonies and intervals. My discussion of the perceptual justification for this principle is based mostly on informal observations rather than specific psychoacoustical principles (**I**: 234 ff.; section 4.1.1 below); nevertheless, psychoacoustical factors clearly underlie these informal observations. The second is the proximity principle of voice leading, whose relationship with streaming is already discussed above (see also section 5.3). The third is rootedness, a property of harmonies deriving from their relationships with the harmonic series. The correspondent psychoacoustical phenomenon is virtual pitch, the pitch percept retrieved from the pattern of harmonics in complex tones (see section 5.2). The fourth is the *proximity principle of spacing*, the avoidance of small intervals (semitones and whole-tones) in consonant harmonies. This principle is related to the phenomenon of critical band; violations of the proximity principle of spacing causes all corresponding partials of the two tones to be within a critical bandwidth with each other, producing effects of roughness and masking (section 5.1).

While the number of these considerations is not great, some of them involve fundamentally important principles. The first two are crucial for all the present examples. In addition, the third, rootedness, is indispensable for explaining the relationship between the bass and upper voices in the analyses in **II** and **III**.

---

[41] This does not mean that perceptual studies could not be relevant to several aspects of prolongational theory. Interesting studies addressing the perception of prolongational structures include Dibben 1999 and Wagner 1990.

The analytical observations in **I–III** relate to perception in widely variable ways. In some cases the relationship is obvious, while in others it is very subtle. This accords with basing the criterion of theoretical and analytical validity on the "musical experience of a sensitive listener," since such an experience involves a multitude of obvious as well as subtle features. While both obvious and subtle features may be significant, they relate differently with prolongational structures. An important function of prolongational syntax is to support and clarify subtle or "concealed" relationships, but such support and clarification presupposes that the functional norms regulating the syntax are based on more obvious features. This principle pertains to the perceptual arguments for the need to use register-sensitive concepts in the treatment of consonance and dissonance (section 4.1.1). It also pertains to the proximity principle of voice leading. As discussed above, the enlargement of time scale tends to make voice-leading connections less obvious for perception, a tendency that has lead some critics of prolongational analysis to altogether deny the experiential significance of large-scale voice leading. While I do not agree with such scholars, I do believe that such a tendency must be counteracted by principles of voice leading that are perceptually as robust as possible—a requirement met by the proximity principle of voice leading.

3. THE CONCEPT OF PROLONGATION

Subsection 3.1 briefly considers the concept of prolongation on the basis of the Schenkerian view of conventional tonality. It also discusses the relationships between prolongation and concepts based on temporality and centricity. The purpose of this discussion is not to treat Schenkerian theory in a formal or comprehensive way but to offer background for the subsequent considerations on post-tonal prolongation. This means, on the one hand, pointing out features crucial to prolongation in both tonal and post-tonal music and, on the other, identifying some aspects that distinguish post-tonal from tonal organization without contradicting the crucial requirements of prolongation. Such requirements will then be treated in a more comprehensive way in section 3.2 on the basis of Straus's conditions.

3.1 PROLONGATION, TEMPORALITY, AND CENTRICITY

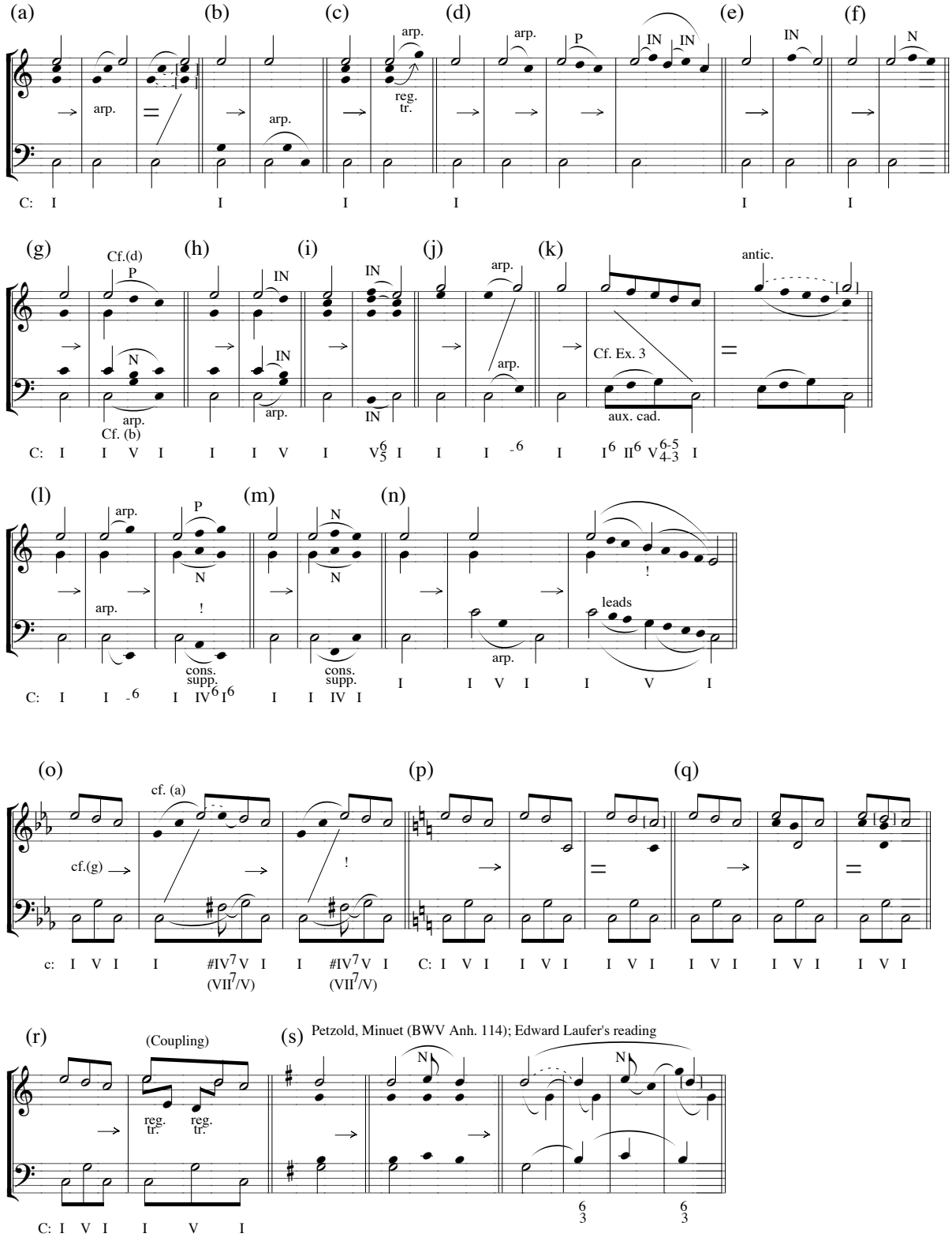
3.1.1 PROLONGATION AS A STRUCTURAL RELATIONSHIP

Schenker and his followers have used the term *prolongation* in somewhat divergent ways. In the present studies, the term is used for all relationships between "higher" (or "deeper") and "lower" levels in Schenkerian or quasi-Schenkerian structures.[42] Example 5 shows examples of prolongation in conventional tonality using conventional graphic notation. The element or progression to the left of an arrow (the higher level) is prolonged by the elements to the right of it (the lower level).

The basic experiential relevance of prolongation is that the lower level may be perceived as an elaborated version of the higher level. The most important category of prolongational operations is *embellishment* (or diminution): one element at a higher level (chord, interval, or

[42] Hence, in the present usage, the *Ursatz* prolongs the triad, and, at the other end of structural levels, surface figuration prolongs unfigured elements. To my awareness, Schenker did not use the term in either of these meanings, but preferred *Auskomponierung* (composing-out) as a general term. For example, Schenker 1935/1979 (§45 ff.) identifies prolongation as a designation for the voice-leading levels at the middleground level. On the other hand, he also spoke of prolongation of *rules*; i.e., the application of simple principles to more complicated cases. The present usage, in which prolongation is a general term for relationships between "higher" and "lower" structural levels, corresponds to that of several more recent Schenkerians and is adopted for practical reasons. It also corresponds to the notion of prolongation manifest in Straus's conditions.

EXAMPLE 5. Prolongational operations in conventional tonality



tone) is substituted by several consecutive elements at a lower level, according to certain syntactic principles (Example 5a–n). Other prolongational operations include temporal displacement (Example 5o, exclamation mark) and registral displacement (Example 5p–q).[43]

A crucially important feature of the first category (embellishment or diminution) is its potential to create well-defined hierarchical structures. The elements on lower levels may be prolonged by further embellishments, creating further structural levels. The highest hierarchical level consists of the tonic triad—replaced by some other referential harmony in post-tonal circumstances—which is prolonged by the elements of the next level (the *Ursatz*; see the outer voices in Example 5g, or the more customary Schenkerian notation in Example 5o, ff.).[44] These elements, in turn, are prolonged at the next level, and so forth up to the lowest level, which contains all the notes in the actual music.

The simplest form of embellishment occurs in one voice and involves no chord progression (Example 5a–f). There are two basic types of embellishment. The first of them is arpeggiation, the successive presentation of harmonic tones (Example 5a–b). Regarding arpeggiation, we should distinguish between two aspects of prolongation. On the one hand, the tones of an arpeggiation prolong the arpeggiated *harmony* (chord) or a part of it, by transforming a simultaneity to a temporal succession of tones. On the other hand, they prolong a *tone* at either end of the arpeggiation, in a structurally superior voice. The latter aspect involves the crucially important principle that voices are not of equal structural status: two of them, the structural outer voices, are superior. Hence in Example 5a, the arpeggiation G4–C4–E5

---

[43] These categories may be compared with Larson's (1997: 119–20) "list of transformations that may be heard as creating embellishment figures [Larson equates embellishment with prolongation]: (1) the addition of an affix or connective, (2) the registral shifting of a note, (3) the temporal shifting of a note, (4) the elision or overlap of shared pitches, and (5) the suppression of an implied tone." In this list, (1) represents embellishment in the present sense; (2) and (3) are also among the present categories. (4) and (5) have less far-reaching significance for prolongation. An example of (4), the elision or overlap of shared pitches, may be seen in the present Example 5o, in which the E♭5 in the last level corresponds to two E♭s in the preceding level. (5), the suppression of an implied tone, is not understood here as a relationship between structural levels but between the "lowest" structural level and the actual music.

[44] To be precise, if we conceive of the formation of the *Ursatz* as illustrated in Examples 5b, 5d, and 5g, there is a structural level between the triad and the *Ursatz*, which consists of the top-voice arpeggiation $\hat{3}$ –$\hat{1}$ (Example 5d). According to this conception, the primary tone (*Kopfton*) is structurally superior within the *Urlinie*. To my awareness, Schenker did not make unequivocal assertions concerning the structural order between the first and last tone in the *Urlinie*; his standard notation (Example 5o) is impartial in this respect. In the present conception, the *Urlinie* is interpreted on similar grounds as any other third-progression. The superiority of the *Kopfton* is supported by its registral height and by its coming first; *Ursatz*-like structures at "lower" levels most often prolong the opening top-voice tone and there is no justification for making the reverse interpretation at the background level. Incidentally, an opposite view to the present one is held by Lerdahl and Jackendoff (1983: 179 ff.), who consistently interpret the concluding $\hat{1}$ as structurally superior, owing to their conception of prolongation as based on tension and relaxation. However, while there is certainly an aspect in which the octave is more "consonant" or "relaxed" than the third, this aspect does not consistently correlate with prolongational weight: outer-voice octaves, in general, are not favored in relation to thirds in any plausible conception of prolongation in conventional tonality. Hence, while effects of tension and relaxation (or stability and instability) are certainly relevant for prolongation, it would seem that Lerdahl and Jackendoff's application of such concepts is overly indiscriminate. Another less cogent feature in Lerdahl and Jackendoff's interpretation is the structural conflict it involves between the outer voices of the *Ursatz*: the superior element is at the conclusion in the top voice but at the beginning in the bass (bid: 275). It is not easy to see how such a conflict would reflect musical experience.

prolongs the C major chord, on the one hand, and the top-voice E, on the other.[45] (The structural order of voices is reflected in analytical notation, which often shows only the structural outer voices, adding inner voices when relevant to illustration.)

In Example 5a–b, arpeggiation connects tones that are present at the higher level in the same register. In Example 5c, arpeggiation adds a tone, G5, which belongs to the C major harmony, but is not present in that register at the higher level. The G5 may be understood as being produced by a registral transfer. Such registral transfers occur frequently in arpeggiation and should be kept distinct from registral displacements of the kind illustrated in Example 5p–q, in which a tone at a higher level is simply replaced by an octave-related tone at a lower level.

The second basic type of embellishment involves non-harmonic tones, which connect stepwise with harmonic ones.[46] Types of non-harmonic tones include passing tones (Example 5c), neighboring tones (Example 5d), and incomplete neighbors (Examples 5c and 5e).[47]

More elaborate prolongation combines motion in several voices, producing chord progressions (Example 5f–m). In such progressions, all individual voices may consist of embellishment figures within the governing harmony, as in Example 5f–j. However, there are also progressions in which only one of the structural outer voices forms such figures and the other one counterpoints it more freely. The bass, in particular, often includes leaping intervals which do *not* arpeggiate the governing harmony but freely counterpoint a stepwise progression in an upper voice, producing consonant support for the passing or neighboring tone (Example 5l–m). In the top voice, such non-arpeggiating leaps are less characteristic at the surface; however, they may form the framework of stepwise progressions against a *leading*, or structurally determinative, bass, as in Example 5n (cf. Schenker 1935/1979: § 221, Fig. 95e3).

As a norm, the structural outer voices coincide with each other with respect to temporal pacing and structural values, but there are several kinds of exceptions to this norm, as shown by the slanted lines in Example 5. The arpeggiation in Example 5a offers a very simple example: the top-voice E appears later than the bass C. Example 5o is more complex since the temporal displacement causes the goal of the top-voice ascent (E♭) to appear above another harmony. Examples 5j–k show another kind of non-correspondence between the outer voices. While they coincide with each other with respect to temporal pacing, the structurally superior tones of each voice occur at the opposite ends of the temporal span.

---

[45] Regarding the concept of "voice," one should bear in mind the distinction between structural levels. It may seem contradictory that arpeggiation has been identified as embellishment in one voice, since it is also frequently described as motion from a voice to another. This contradiction vanishes by specifying that "voice" refers to a lower level in the former expression and to a higher level in the latter.

[46] The distinction between harmonic and non-harmonic tones cannot be made purely on the basis of the membership in the governing harmony. For example, in a seventh chord, an octave may function as a neighbor of the seventh, and, thus, somewhat paradoxically, as a non-harmonic tone. Such "paradoxical" situations are characteristic of some of the non-triadic music discussed in the present studies; see the discussion of *h-neighbors* and *h-passing tones* in section 6.2 and in **II**.

[47] Of other significant types of non-harmonic tones, suspensions may be technically defined as prepared incomplete neighbors and cause no special problems for the present considerations.

As already observed, the "lowest" level in the prolongational structure contains all the notes in the actual music. In addition, it may contain *implied* tones, tones that do not actually occur. In analytical graphs, such tones are shown in parentheses or brackets; the present studies use square brackets. A very simple instance of implied tones is shown after "=" in Example 5a; it involves the mental presence of tones after their physical presence has ceased. Registral displacements may also be understood in terms of implied tones. They imply the presence of a tone in a register in which it does not actually occur; see the alternative notation after "=" in Example 5p–q. However, there are also instances of implied tones that do not appear at any time or any register, as in Example 5s, which depicts the opening of Christian Petzold's well-known Minuet (J. S. Bach's *Notebook for Anna Magdalena Bach*, BWV Anh. 114).[48] In this excerpt, the D in m. 4 does not occur in any octave.

The justification required by readings based on implied tones is extensively discussed in Rothstein 1991. As explained in this study, such justification involves both melodic-contrapuntal and harmonic principles. Melodically, the bracketed notes in Examples 5p–q and 5s complete "archetypal" patterns, simple passing or neighboring figures. Harmonically, these implied tones rely on the notion of "imaginary continuo." This notion involves the implicit presence of chord tones in all registers, enabling the readings based on registral displacements. It also involves the supplementation of incomplete harmonies to form triads; for example, while the last chord in Example 5s only includes B and G, the D is present in the 6/3 chord of the "imaginary continuo."

Above, I have discussed the principles of prolongational relationships in conventional tonality in a most general manner. The more specific techniques, as discussed by Schenker, are derivable from these general principles (for example, *unfolding* is formed by specific combinations of arpeggiations, sometimes connected with temporal displacements [Schenker 1935/1979: § 140 ff., Fig. 43]). Turning now to prolongation in post-tonal music as evident in the examples of the present studies, it may be noted that most of these general principles are applicable to these examples by replacing the triad by other kinds of referential harmonies (which vary from case to case). Hierarchical levels are formed by elaborating the referential harmony with embellishments consisting of either arpeggiation or "stepwise" motion that forms passing and complete or incomplete neighboring figures.[49] The present post-tonal examples also observe the principle of the structural differentiation of voices; most of them manifest the predominance of the structural outer voices. However, there are also some noteworthy differences between the present approach and conventional Schenkerian theory,

---

[48] Example 5s is based on a lecture handout of Edward Laufer (Helsinki, 1992).

[49] "Stepwise" is in quotes when it refers semitones and whole-tones in post-tonal music, because the distinction between "steps" and "leaps" cannot (usually) be based on a referential collection equivalent to the diatonic set in tonality, but relies on the absolute size of the interval. Moreover, by "sevenths" and "ninths" (in quotes), I shall refer to larger realizations of interval classes 1 and 2 regardless of enharmonic notational variants (for example, a "seventh" may be notated as an augmented sixth or a diminished octave).

reflecting differences between post-tonal and tonal music. The present approach sets no premises equivalent to the *Ursatz* forms with respect to how the referential sonority is elaborated on the background.[50] The background structures are determined purely by case-specific analyses. The reasons for this difference are easy to understand. As discussed in section 2.1, the principles of organization are much more piece-specific in post-tonal music and less liable to generalizations such as the *Ursatz*. (Moreover, since our present knowledge of post-tonal prolongation is much more narrow and rudimentary, it offers less grounds for making generalizations.)

Another difference between conventional Schenkerianism and the present approach to post-tonal music is the lesser importance of implied tones in the latter. To be sure, the simplest type of implied tone in which the mental presence of a tone is longer than its physical presence, as in Example 5a, occurs frequently in the analyses of **I–III**. Such simple instances are based on our general capacity of memory rather than any style-specific circumstances. However, registral displacements in which "sevenths" or "ninths" substitute for "stepwise" intervals in voice leading (as in Examples 5p–q) are less common in the present analyses; insofar as they occur, they require greater contextual justification. Moreover, suppressed tones with no representative in register (as in Example 5s) do not occur at all in the present analyses of post-tonal music. Harmony and voice leading are not based on "imaginary continuo" but depend more strictly on the concrete presence of notes.

This difference between tonal and post-tonal organization may be seen as being related with a difference in the strength in which the norms of harmony and voice leading are established in the listener's cognition. While both kinds of analysis assume that such norms, largely deriving from a referential harmony, are established in an experientially relevant way, it seems reasonable to assume that such establishment is more robust in the case of conventional tonality and thus more readily supports the cognitive supplementation of incomplete harmonies. Two kinds of considerations support this assumption. First, the establishment of the norms of conventional tonality is supported by the extensive "common practice" repertoire, whereas the norms of post-tonal organization apply to more narrow bodies of music, such as individual works. Second, the adoption of conventional triadic norms is also facilitated by the high degree of correspondence that such norms have with perceptual phenomena. These considerations relate with the model in Example 1: conventional triadic principles are strongly supported both by the size of the general context and by psychoacoustics. However, it should be added that there is no completely straightforward way in which psychoacoustical principles favor triads over other kinds of referential harmonies; some of the present examples (in **II**, especially) also show a fairly consistent correspondence between the constitution of the referential harmony and

[50] The customary *Ursatz* forms are not the only possible ways to prolong the triad. Whether there are background structures alternative to the *Ursatz* forms in triadic music is an issue outside the present considerations.

the psychoacoustical properties of intervals. As a more straightforward factor one may mention simply the number of tones, which is smaller in triads than in any of the referential harmonies of the present examples, facilitating the cognitive task of complementing harmonies by implied tones. Among three-note chords, again, the interval content of triads shows a remarkably high correspondence with the psychoacoustical background factors of consonance (see section 5.2.3).

The present studies also point out a specific reason why "stepwise" intervals are less readily replaced by "sevenths" and "ninths" in the voice leading in post-tonal prolongational structures: unlike triads, the referential harmonies in the present post-tonal examples include "sevenths" and "ninths" as consonant intervals. Therefore the use of such intervals as substitutions for "stepwise" intervals in voice leading creates a more acute danger of confusion between arpeggiation and "stepwise" voice leading, a point to be discussed in connection with Straus's condition #4 (see section 3.2). To avoid misunderstandings, one should observe the difference between registral displacements, in which a higher-level tone is simply replaced by an octave-related tone (as in Example 5p–q) at a lower level, and registral transfers, in which a note occurs in both the "due" register and other registers (as in Example 5r). The latter kind of prolongation occurs frequently in the present post-tonal examples. For example, in the structure of *Voiles*, as shown in Example 4b–c above, the D7 is prolonged by registral transfers, but there is no registral displacement comparable to Example 5p–q since D7 is also present and connects "stepwise" to D♭7.[51]

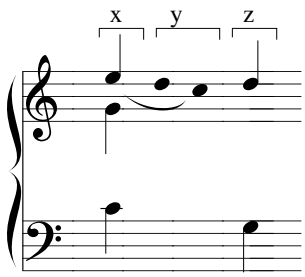
3.1.2 PROLONGATION AND TEMPORALITY

According to the present definition, prolongation is a *structural* relationship between different hierarchical levels and between elements at such levels: the "lower" levels prolong the "higher" ones. However, in ordinary language, "prolongation" also has a *temporal* meaning, which may coincide to a greater or lesser degree with the structural meaning. For example, consider the way in which Straus opens his discussion of prolongation. In reference to a simple example representing conventional tonality, reproduced here as Example 6 (Straus 1987: Example 1), he explains: "Given three musical events X, Y, and Z, like those in Example 1 [the present Example 6], the prolongational model claims: 'Y is structurally inferior to X and extends X; X is not displaced until Z arrives'." (Ibid.: 2.) This explanation combines structural relationships with temporal ones: "Y is structurally inferior to X" but also "extends X" temporally. It seems natural to say that Y prolongs X in both the structural and the temporal

[51] Hence in the present usage registral displacement means a relationship between structural levels but registral transfer means a temporal relationship within a structural level. Registral transfer may technically be defined as a special kind of arpeggiation.

sense. A similar correspondence between prolongation as a structural relationship and in the temporal sense is evident in the *Ursatz* progression, as shown in Example 5g (in which the X and Y of Example 6 are combined with bass arpeggiation). However, while such progressions seem archetypal examples of prolongation, prolongation as a structural relationship and temporal prolongation do not always coincide in an equally straightforward manner. In order to avoid confusion, it is worthwhile to make a closer look at the temporal relationships in prolongational structures.

EXAMPLE 6. A reproduction of Straus 1987: Example 1



Two features are significant for the observed correspondence between the structural and temporal meanings of prolongation in the progressions in Examples 5g and 6. First, the prolonged elements—the tonic harmony and the structurally superior voices—appear at the beginning of the progression. Second, the end of the progression employs the same harmony as the beginning; consequently, even though the top voice has moved from $\hat{3}$ to $\hat{1}$, we can easily imagine or auralize the starting point as being present at the end point. These two features are not shared by all prolongational structures. First, prolonged elements (in the structural sense) may appear only at the end of a progression. For top voices, this is exemplified by opening ascents (stepwise or arpeggiating) to the *Kopfton* (Example 5o); for basses and harmonies, by auxiliary cadences (Example 5k; cf. Schenker 1935/1979: § 244–5). Second, even if the prolonged entities appear at the outset, the end of the progression may be occupied by another harmony, such as a back-relating dominant (Example 5h; see the discussion of *applied dividers* in Schenker 1935/1979: § 279). In such cases, we cannot auralize the starting point as being present at the end point, and the temporal effect of the latter might be described as "interrupting" the presence of the structurally superior elements, rather than "prolonging" them.[52]

[52] Such a description of "interrupting" temporal effects may invoke the concept of "interruption" in the technical sense (Schenker 1935/1979: § 87 ff.), which, indeed, is not unrelated with progressions such as Example 5h. Schenker (ibid.: § 279) bases his discussion of applied dividers, such as the V here, on the transference of the interrupted structure, contending that "[t]he characteristic feature of the interruption must always be present." He does not identify concomitant upper-voice motions in terms of incomplete neighbors as is done in Example 5h, but such a back-relating identification would seem logical in view of the accompanying harmonic relationship. As regards actual interrupted structures, Schenker (ibid.: § 91) emphasizes that "[t]he first

While one should thus be wary of equating prolongation as a structural relationship with the ordinary temporal meaning of "prolongation," temporal circumstances are certainly relevant to the structural relationship. The structurally superordinate elements normally occur at either end, or both, of the temporal span in which they are prolonged. Given a structural element, S, and an elaborative element (or a chain of elements), E, there are thus three schemata for temporal successions prolonging S, namely, SES, SE, and ES.[53] If S and E are harmonies, the above discussion implies that the SES schema corresponds best to the temporal meaning of prolongation. This schema occurs in the normative I–V–I progression in tonality, whereas SE (I–V) and ES (V–I) are "incomplete" versions of this progression. If S and E are tones within one voice, a similar distinction applies to complete (SES) and incomplete neighbors (SE or ES). For top-voice arpeggiations—with or without passing tones—the SE schema fits well the temporal meaning, but only when accompanied by an SES harmonic progression (or a stationary S). All this may seem to suggest that SES, a departure-and-return schema, is a prerequisite for the most "complete" form of prolongation. However, even though the normative background structure in conventional tonality requires this schema—exceptions to this norm are given by pieces with the auxiliary cadence as the background structure (ES)[54]—one should be cautious about generalizing this requirement to post-tonal circumstances.

Also Straus (1987, 6) emphasizes that a clear distinction should be maintained between prolongation and departure-and-return schemata. While these two organizational features coexist in the *Ursatz*, departure-and-return is neither a sufficient nor a necessary condition for prolongation. It is not sufficient, because the juxtaposition of arbitrary elements A, B, and A does not guarantee that the relationships between A and B can be described in terms of prolongation. It is not necessary, because the SE and ES schemata exist. That we are able to distinguish between these two schemata makes evident that the structural order in prolongational syntax cannot be derived from temporal order; Examples 5h (SE) and 5i (ES) illustrate that the structural order depends on pitch-based norms. In the tonal I–V–I progression, the primacy of I is supported both by such pitch-based norms and by temporal relationships, but one is not justified in supposing that these aspects always coincide in post-tonal prolongation even at the highest level (this discussion is pursued in section 3.2.3.5).

$\hat{2}$ is not a neighboring note" but a passing note; however, the passing function presupposes the eventual arrival of $\hat{1}$, which does not take place in the case of applied dividers. (Moreover, it has often been observed that Schenker is not consistent regarding the question as to whether the first or second $\hat{2}$ is structurally decisive in interrupted structures, an issue is too far removed from my main concerns to be dwelt on here.)

[53] ESE does not have equal significance and is left out of the present considerations. In neither conventional tonality nor in the present examples does this schema have any independent significance in stepwise embellishments or in harmonic progressions; it can be formed only as a combination of ES and SE. Such an exclusion of ESE (or at least bias in favor of SES) would seem to be supported by rather universal *Gestalt* principles. In conventional tonality, ESE has some significance only in upper-voice arpeggiations, as in unfoldings proceeding from an inner voice to the top voice and back (Schenker 1935/1979: Fig. 43a2).

[54] See, for example, Schenker 1935/1979: Figs. 110a3, 110d2. In addition, the SE schema (I–V) is found in some Baroque preludes (ibid.: Fig 152,6), but its significance for independent compositions is smaller.

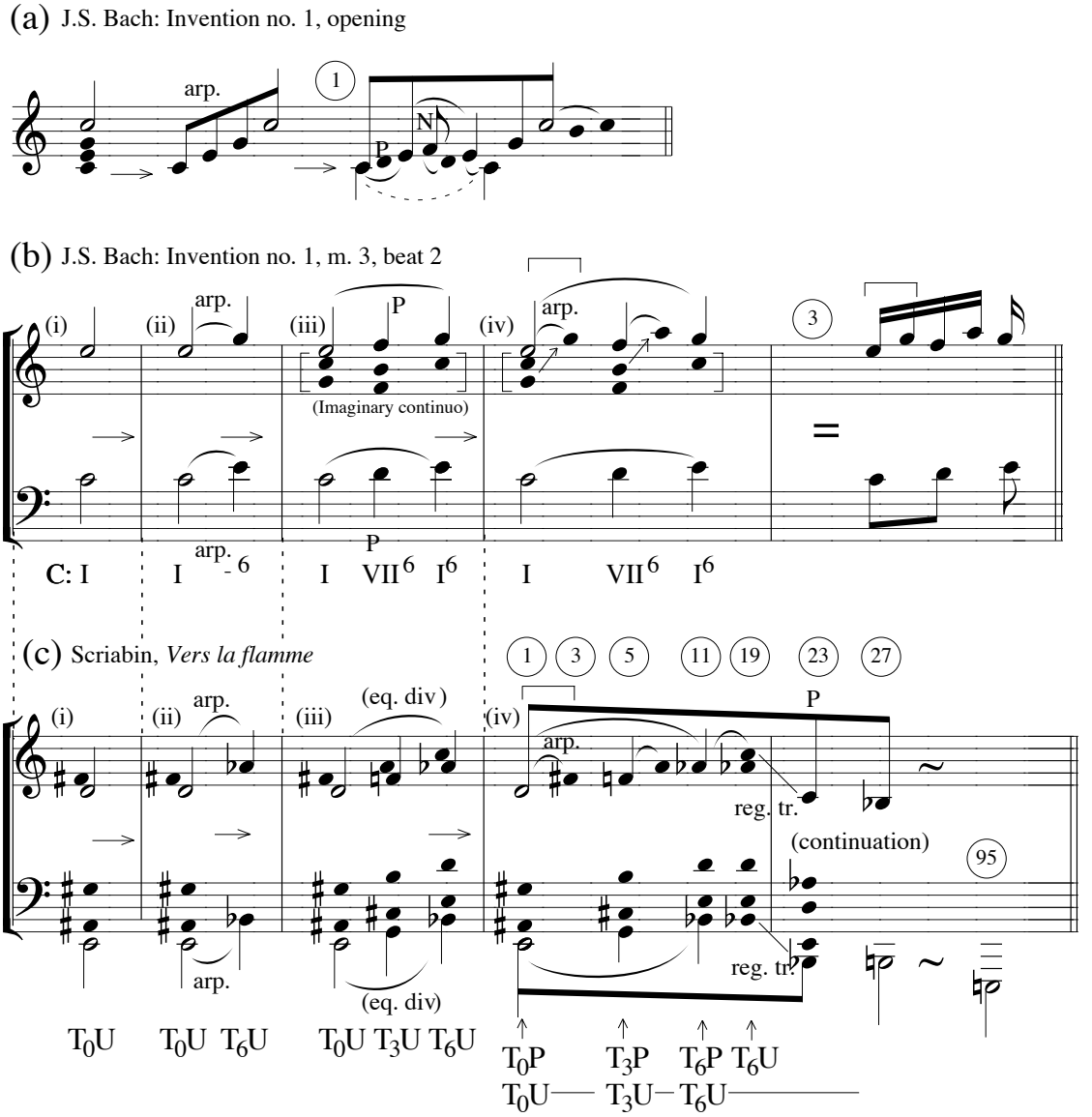
Another kind of complexity within the structural and temporal aspects of organization is evident in cases in which a "higher"-level simultaneity is prolonged by non-simultaneous elements. Such cases include arpeggiations and outer-voice asynchronies (Examples 5a, 5j-k, 5o). The issue of arpeggiation is especially significant for the present considerations, because it involves a noteworthy difference between tonal and post-tonal organization, stemming from the lack of "imaginary continuo" in the latter. Consider an unaccompanied arpeggiation or, more generally, one proceeding to a tone or tones that are not initially stated in any register (permitted by the prevalent conception of harmony; see section 4.1.2). Such an arpeggiation prolongs structurally the harmony it arpeggiates, but its temporal effect may rather be described as the "enlargement" or "complementation" of harmony. However, owing to the presence of the "imaginary continuo," such an effect is only apparent in tonal music and has less important structural ramifications. For example, in the opening arpeggiation of Bach's Invention no. 1 (Example 7a), the G completes the C major triad, which makes a satisfactory aesthetic effect but is not indispensable for establishing the identity of the harmony. Even prior to its literal occurrence, G is understood as an implied tone. In post-tonal circumstances, such enlargement or complementation must be taken at its face value; incomplete forms of harmonies are not mentally complemented prior to the literal complementation. The present studies contain several examples in which the enlargement of harmony through arpeggiation becomes an important structural feature and involves considerable temporal dimensions.[55]

One such example is Scriabin's *Vers la flamme*, whose opening section (mm. 1–26) is depicted in Example 7c (adapted from **II**: Example 8). This example is vertically aligned with another excerpt from the Bach Invention (Example 7b), in order to demonstrate a structural analogy between these passages; this analogy will be taken up after discussing the role of foreground arpeggiations in the Scriabin. The opening harmony (mm. 1–4), designated as T_0U, consists of E, A♯, G♯, D, and F♯, but the opening block chord is its subset T_0P, which omits F♯, major ninth in relation to the bass (fb-interval 2 in the terminology used in **II**; see section 4.1.2). F♯ appears in m. 3, through the arpeggiation D–F♯, but T_0U does not yet appear as a verticality because of concurrent inner-voice motions (not shown in Example 7c) and the absence of D. Whereas in the Bach example of Example 7a—or in the bracketed sixteenth-note figure in 7b—the fifth (G) is implicitly present in the harmony before its literal occurrence, there is less justification for reading an implicit major ninth (F♯) in mm. 1–2 of the Scriabin passage.

---

[55] For a point of comparison, one may consider processes in tonal music which postpone the establishment of the primary top-voice tone (initial ascents), the primary bass note (auxiliary cadences), or the tonic level (ditto). In the present post-tonal examples, comparable postponement may also concern the construction of the referential harmony in a way foreign to conventional tonality.

EXAMPLE 7. Examples of arpeggiation in tonal and post-tonal music



The difference between chord U and its subset P also bear more important structural ramifications than any difference between a complete and incomplete triad in tonal music. Interpreting the major ninth (F♯) as part of the opening harmony (T_0U) relies on the proximity principle of voice leading (see sections 2.2.1 and 3.2.3.4), under which the "leap" D–F♯ constitutes arpeggiation. This is later concretized—in the sense discussed in section 2.2.1—by the block-chord T_6U in m. 19, after two sequential repetitions of the opening material. This concretized major ninth also assumes a significant structural function as a middleground passing tone (after a downward registral transfer; see Example 7c, m. 23, C4).

The Scriabin passage illustrates a tendency characteristic of the examples of the present studies. Since referential harmonies are more complex than in triadic tonality and since their

members are not felt as implicitly present, their internal relationships offer more material for musical elaboration. Such relationships are often exploited in the presentation of individual harmonies, for example, by focusing on special subsets and their complementation. Consequently, individual harmonies often occupy larger temporal spans, rendering the rhythm of harmonic change slower than what is typical of conventional tonality. The three harmonies in the 26 measures of the Scriabin passage (T_0U, T_3U, and T_6U) are not prolonged through subordinate chord progressions, as would be the case in typical tonal passages of comparable length. (The enlargement of P to U has the nature of surface figuration rather than actual chord progression.)

A more extreme example of such a tendency is given by Debussy's *Voiles* (Example 4). All the whole-tone material, which constitutes the greatest part of the piece, may be interpreted as figuration within a single harmony. The internal relationships and oppositions within the harmony thus become decisive in the structure. Moreover, whereas in the Scriabin it takes three measures to complement chord P to form U, in the Debussy the referential harmony is complemented through an arpeggiation that reaches its goal (D7) only in m. 22.[56] (As discussed in **II**, this ties in with an especially sharp dichotomy between prolongation in the structural sense and temporal connotations of prolongation.)

I shall return to these unconventional aspects in the temporal presentation of harmonies in section 6.1.2. At present, it may be worthwhile to balance these observations by pointing out that otherwise the structures in tonal and post-tonal prolongation may be closely analogous, as illustrated by the vertical alignment of Examples 7b–c. Graphs (i)–(ii) indicate that the structure is in both cases based on outer-voice arpeggiation within the governing harmony. In the Bach, the top-voice E5–G5 and bass C4–E4 form parallel tenths and arpeggiate, of course, the C major triad. In the Scriabin, the top-voice D4–A♭4 and bass E2–B♭2 form parallel minor sevenths—a consonance in this context—and are arpeggiations within T_0U (more precisely, they arpeggiate the two superimposed tritones in the opening T_0P, another manifestation of the special significance of this subset). Graph (iii) shows the elaboration of these arpeggiations by passing tones in the Bach and by an equal division of the tritones in the Scriabin.[57] Finally, graph (iv) shows similar surface arpeggiations between the two excerpts, although, as discussed above, the structural significance of such arpeggiations is different in each case. In the Bach, the bracketed E–G arpeggiation is less essential for establishing the presence of G in the harmony since it is present anyway in the "imaginary continuo."[58]

---

[56] *Voiles* and the *Vers la flamme* passage share chord U as the referential harmony (understood in terms of bass-related intervals). However, these examples involve different internal distinctions within U; compare the opening subsets Q and P in Examples 4b and 7c, respectively.

[57] This equal division may be interpreted as passing-tone motion by stretching the region of "steps" in the application of the proximity principle of voice leading (**II**).

[58] Whereas this G5 may be interpreted as a registrally transferred inner voice, a similar interpretation is not available for the F♯4 in m. 3 of *Vers la flamme*, because there is no imaginary continuo to contain it as an inner voice in mm. 1–2. Moreover, the spacing in which the major ninth lies above the minor seventh is generally

3.1.3 PROLONGATION AND CENTRICITY

Another distinction that Straus (1987) emphasizes is that between prolongation and centricity. There are several ways to promote to centric effects, many of which have nothing to do with the syntactic norms required by prolongation.[59] Centricity is thus clearly not a sufficient condition for prolongation; whether it is a necessary condition is a somewhat more complicated question. Prolongation may be regarded as one of the factors relevant to centricity: prolonged harmonies are centers in a certain sense. In tonal compositions, clues to the center given by prolongational and non-prolongational aspects usually agree, but this is not necessarily the case in post-tonal music. As discussed in **I** and in section 3.2.3.5 below, Schoenberg's op. 19/2 is a case in point. According to my analysis, the *opening* harmony is superordinate in the prolongational structure, because the overall structure may be understood as its embellishment. However, rhetorical, registral, and textural aspects impart a centric quality to the *closing* harmony. This example also illustrates that temporal relationships tend to play somewhat different roles for prolongation than for centricity. As discussed above, prolongational superiority may coincide with either the beginning or the end of a temporal span (but the temporal connotations of prolongation fit best with the former possibility). On the other hand, centricity, as it is normally understood, would seem to be goal-oriented rather than departure-oriented.

If one understands *center* as referring to a harmony, the choice of that harmony may thus be less unequivocal in post-tonal music even when the requirements of prolongation are fulfilled. Moreover, if one understands center as referring to a pitch class, the choice of the centric pitch class within the referential harmony may be less unequivocal in post-tonal music, even when both prolongational and non-prolongational factors of centricity favor the same harmony. Consider the several functions concurrently fulfilled by the tonic pitch class in conventional tonality. First, it is the bass and the root of the prolonged harmony. Second, it is the bass and the root of the framing harmony in the temporal organization (or, in pieces based on the auxiliary cadence, the goal harmony). Third, it is the goal of the top-voice progression (*Urlinie*)—which, again, exemplifies centricity as a goal-oriented aspect of organization. The coincidence of these three functions promotes to the status of the tonic pitch class as the unequivocal center, but there is no reason why these functions need coincide. In fact, they do not coincide in any of the post-tonal examples discussed in **I**–**III**, a tendency which makes the choice of a centric pitch class much less unequivocal.

---

characteristic of *Vers la flamme*. Hence the F♯4 must be understood as appearing in its "due" register above the structural top voice D4. The structural priority of the latter is not based on its highest registral position even in a conceptual sense, but on other emphasizing factors.

[59] For example, Luciano Berio's *Sequenza* VII for oboe is clearly centered on B, but it is highly questionable whether its organization manifests prolongation in the Strausian sense.

3.2 STRAUS'S CONDITIONS

3.2.1 INTRODUCTION

Above, we have discussed prolongation on the basis of tonal examples and observed the dependence of prolongational structures on pitch-based norms. Straus's four conditions of prolongation specify the requirements for such norms, without tying them to the particular features of conventional triadic tonality. Here is how Straus identifies the conditions (ibid.: 4–5; italics original):

> "Condition #1. The consonance-dissonance condition: A consistent, pitch-defined basis for determining relative structural weight. […] Condition #2. The scale-degree condition: A consistent hierarchy of consonant harmonies. […] Condition #3. The embellishment condition: A consistent set of relationships between the tones of lesser and greater structural weight. […] Condition #4. The harmony/voice leading condition: A clear distinction between the vertical and horizontal dimensions."

One may conceive of these conditions as addressing two basic aspects of pitch-based functions. Conditions #1 and #2 concern the functional stability of harmonies, conditions #3 and #4 concern functions of melodic motions. The connection of each condition with the harmonic and melodic realm is made evident by the above definitions, except for condition #4, which requires some clarification. From Straus's discussion it is evident that the essential practical requirement is that "we need to be able to distinguish motions within a voice from motions between voices." (Ibid.: 7.) In other words, we need to be able to distinguish melodic (horizontal) intervals that stand for vertical ones at some structural level from those that are purely horizontal connectives, i.e., occur in passing and neighboring figures.[60] The former are, of course, arpeggiations; the latter (seconds in tonal music) are called *voice-leading intervals* in both Straus 1987 and in the present studies.

3.2.2 THE CRITIQUE OF THE CONDITIONS

Before discussing these conditions in detail, let us briefly consider the critique addressed against these conditions (see also **I**: 243–4). In this critique, two lines of argument may be distinguished. According to the first, the conditions are not met even by conventional tonality and thus cannot be a measure of prolongation (Larson 1997). According to the second, the conditions are tailored to fit only conventional tonality; hence it is circular reasoning to conclude

[60] In principle, vertical intervals might also stand for horizontal connectives; a simple traditional example is given by *acciaccaturas* that are struck concurrently with the main note. However, such a phenomenon plays no role in the present analytical examples and may be left out of consideration.

that non-tonal music is not prolongational if it does not fulfill such conditions (Lerdahl 1989, Travis 1990).

The stand of the present studies is that both of these—seemingly contradictory—arguments have only limited validity. As to the first of them, tonal music does contain details in which some of the conditions fail (see, for example, **I**: Example 11). However, such details are subordinate to structurally prior elements that do fulfill the conditions. Moreover, even regarding details, an exception to one of the conditions does not indicate its invalidity but rather its "rivalry" with other factors. Such "rivalry" is often evident between organizational clues given by harmonic stability (conditions #1 and #2), on the one hand, and melodic figures (related to condition #3), on the other (see section 3.2.3.3 below).

To disprove the second argument, the charge of circular reasoning, it suffices to refer to the present analytical examples, in which the conditions are shown to be fulfilled—to a significant degree, at least—in ways that deviate from conventional tonality. It should be added, however, that in some of the examples some of the conditions have to be adapted to some extent. In particular, the system of harmonic stability may be determined in ways that deviate from Straus's description of conditions #1 and #2 in certain respects. Since such deviations do not jeopardize the basic requirement that norms of harmonic stability exist, it seems justified to relax the conditions with respect to this issue (see section 3.2.3.1 below).

With these specifications, Straus's conditions are successful in identifying a certain type of musical syntax, called "prolongation" in the present studies. They have great practical utility in offering a point of departure for discussing purported post-tonal prolongational structures. Whether we could speak of prolongation under quite different conditions (in the manner of Lerdahl 1989) is a semantic question—as Lerdahl (1997: 153) himself observes—and is not relevant to the present studies.

3.2.3 CONSIDERATIONS OF SPECIFIC CONDITIONS[61]

3.2.3.1 CONDITIONS #1 AND #2: FACTORS OF HARMONIC STABILITY

Conditions #1 and #2 determine a two-stage stability system. According to Straus's description, consonance and dissonance distinguishes between structural and non-structural tones on a small scale (condition #1), whereas the scale-degree system determines the mutual hierarchy of the consonant harmonies on a larger scale (condition #2). The relevant factors for conditions #1 and #2 are chord construction and chord transposition, respectively.[62] Chord construction

---

[61] Most of the present discussion repeats and elaborates points made in **I** and **II**.

[62] Actually, Straus does not specify that "scale degree" hierarchies should be based on transpositional levels but speaks of "some kind of hierarchy among the consonant harmonies" (Straus 1987: 7). In the present studies the notion of "scale degree" systems refers to functional stability hierarchies based on transpositional levels. "Scale

distinguishes consonances from dissonances. Transpositional levels of consonances, in turn, determine the scale-degree hierarchy. However, as observed above, there is reason to allow room for alternative ways in which these factors may pertain to harmonic stability.

First of all, chord construction may provide the basis for more nuanced distinctions than the two-alternative consonance–dissonance system. In such a case, we may speak of a system of gradated consonance.[63] Such gradated consonance may, in part, compensate for a scale-degree hierarchy (see, for example, the analysis of Schoenberg's op. 19/2 in **I**). Moreover, the roles of chord construction and chord transposition on different scales of organization may deviate from conventional tonality. In some cases, the *large-scale* organization is based on varying the construction of the referential harmony. The large-scale relationships may be understood in terms of gradated consonance: the variant harmonies are "less consonant" than the referential harmony. However, such variant harmonies may function as *locally referential*, giving rise to *small-scale* "scale degree" systems based on transpositional levels (still closer to the surface, a more or less ordinary consonance–dissonance relationship may become operative). Scriabin's *Vers la flamme* exemplifies such organization (**II**). The T_0U in the introductory section (Example 7c) functions as an "appoggiatura" chord in relation to the primary referential harmony of the overall organization, but is locally prolonged by small-scale "scale degrees" T_3U and T_6U.[64]

All this suggests that instead of adhering rigidly to conditions #1 and #2, as described by Straus, we should posit a more general requirement that norms of harmonic stability exist. Such norms are based on relationships with the referential harmony, but the roles played by chord construction and chord transposition in these relationships vary.

---

degree" is in quotes since in post-tonal music such systems are not normally based on degrees in a scale equivalent to the diatonic scale in conventional tonality. It should be observed that chord transposition in conventional tonality operates primarily within the diatonic set, which brings about the variation in construction between major and minor triads. In the present examples, transposition is an operation within the chromatic set.

[63] The consonance–dissonance system of conventional tonality is also not devoid of such gradation. 5/3 and 6/3 chords are both consonant but the former are more stable. Moreover, dissonances are not equally incapable to support structural tones: certain dissonances such as V^7 and cadential 6/4 support them so frequently that an adequate theoretical description of conventional tonality must acknowledge their privileged position.

[64] In the overall organization of *Vers la flamme* (see **II**), U is subordinate to A^{+9} which is a superset of U in terms of bass-related intervals. These chords also involve a characteristic difference in registration. As shown in Example 7c, the second lowest tone in the U chord forms the tritone with the bass, which is replaced by the fifth in the A^{+9}; this change offers psychoacoustical support for the functional order between these two chords. Another locally referential harmony, intervening between U and A^{+9} in the overall flow of events, is a chordal inversion of A^{+9}. Another example of large-scale organization exploiting varying chord construction is Debussy's *Ce qu'a vu le vent d'ouest*, in which variants of a referential harmony are obtained by half-step shifts (**III**: Example 2; see also the middle stave of Example 23b in the present essay).

3.2.3.2 CONDITIONS #3 AND #4: ASPECTS OF MELODIC RELATIONSHIPS

As noted above, conditions #3 and #4—the embellishment condition and the harmony/voice-leading condition—concern the interpretation of melodic motions. These conditions are closely interrelated. In his summary of the conditions, Straus (1987: 7) describes these conditions as follows: "Third, the embellishment condition; we need a consistent model of voice leading that will enable us, for example, to tell an arpeggiation from a passing note. Fourth, there is the harmony/voice leading condition; we need to be able to distinguish motions within a voice from motions between voices." According to this description, the requirements made by conditions #3 and #4 may seem confusingly similar. That they actually concern different aspects of organization becomes more evident from Straus's initial identification of condition #3 as a requirement for "a consistent set of relationships between the tones of lesser and greater structural weight." Hence condition #3 involves the correspondence between melodic figures and structural order, whereas condition #4 concerns the distinction between arpeggiations and voice-leading intervals, saying nothing of their structural order.

Understood in this way, neither condition follows logically from the other. On the one hand, there might be a "consistent set of relationships" required by condition #3 which does not observe the harmony/voice-leading distinction. On the other hand, there might be such a distinction without a consistent set of relationships "between the tones of lesser and greater structural weight." For a simple example, suppose that the basis for meeting condition #4 is the familiar distinction between "steps" (voice-leading intervals) and "leaps" (arpeggiations). Consider, then, a figure that moves stepwise from pitch x to pitch y and back to x. Under condition #4, this is motion "within one voice" and not arpeggiation. However, we need condition #3 to specify that such a figure is neighboring motion that typically prolongs x as the structural element.

3.2.3.3 CONDITIONS #1 + #2 AGAINST #3: HARMONIC STABILITY VERSUS MELODIC EMBELLISHMENT AS STRUCTURAL DETERMINANTS

The above considerations lead us to important aspects of the embellishment condition that were not recognized in Straus's original discussion (1987). According to that discussion, harmonic stability (conditions #1 + #2) determines structural weight, whereas the consistent set of embellishments (condition #3) regulates the melodic relationships between the tones of greater and lesser structural weight. Actually, however, melodic figures have a more active role in determining structural weight. Owing to the correspondence between melodic figures and structural relationships, melodic figures are, apart from harmonic stability, capable of offering clues to structural weight. This becomes especially obvious in instances in which the clarity of

an embellishment figure is strong enough to induce a structural order that reverses the norms of harmonic stability. For an example in conventional tonality, see Example 8c below, in which the "I" is subordinate to V^4_3 despite being favored by both conditions #1 and #2 (for similar examples, see **I**: Ex. 11, Larson 1997: Exs. 4–5, Cinnamon 1993: Fig. 1).

The roles of harmonic stability and melodic embellishments as determinants of structural weight are focused on in the discussion between Larson (1997) and Straus (1997b). Larson goes to the extremes, it would seem, in downplaying the former factor: "I have argued that prolongation *is* embellishment; embellishment (and only embellishment) determines the relationships between tones that make some tones of lesser and greater structural weight than others." (Larson 1997: 130; italics original.) Straus, in his response, presents a more balanced view, admitting that conditions #1 and #2 do not determine structural weight in the rigid way that his original description implied, but defending, nevertheless, the significance of these conditions. This view agrees with the present approach. Within a structural level, the structural order is determined by the combination of clues given by harmonic stability (conditions #1 and #2) and melodic embellishments (condition #3). When the former clues give in to the latter, a relatively unstable harmony may be prolonged at that level. Significantly, however, the validity of conditions #1 and #2 is vindicated by the requirement that the unstable harmony must be subordinate to a stable one at a "higher" level (cf. Straus 1997b: 137–8). Hence, the passage of Example 8c requires a larger context, such as that shown in Example 8e, which renders the prolonged V^4_3 (m. 3–4) subordinate to I in the overall progression.

Both kinds of structural clues, "harmonic stability clues" and "embellishment clues" involve gradation: such clues may be more or less strong. The issue of gradated harmonic stability was considered in section 3.2.3.1. The issue of "embellishment clues" will not be treated exhaustively here, nor will the very complicated issue of how the relative weight of the two kinds of clues should be determined in different circumstances. However, some observations of embellishment clues in conventional tonality are useful in shedding light on the analytical decisions in the present studies, which involve largely similar considerations. These observations are by no means original; they represent common knowledge of anyone acquainted with tonal prolongational structures. For one example, most of the following points are manifest in the "preference rules" of Lerdahl and Jackendoff 1983.[65]

First of all, we should observe the significance of meter, rhythm, articulation, and grouping. Metrically or rhythmically emphasized tones are more probably structural, as are the first and last note in a group.

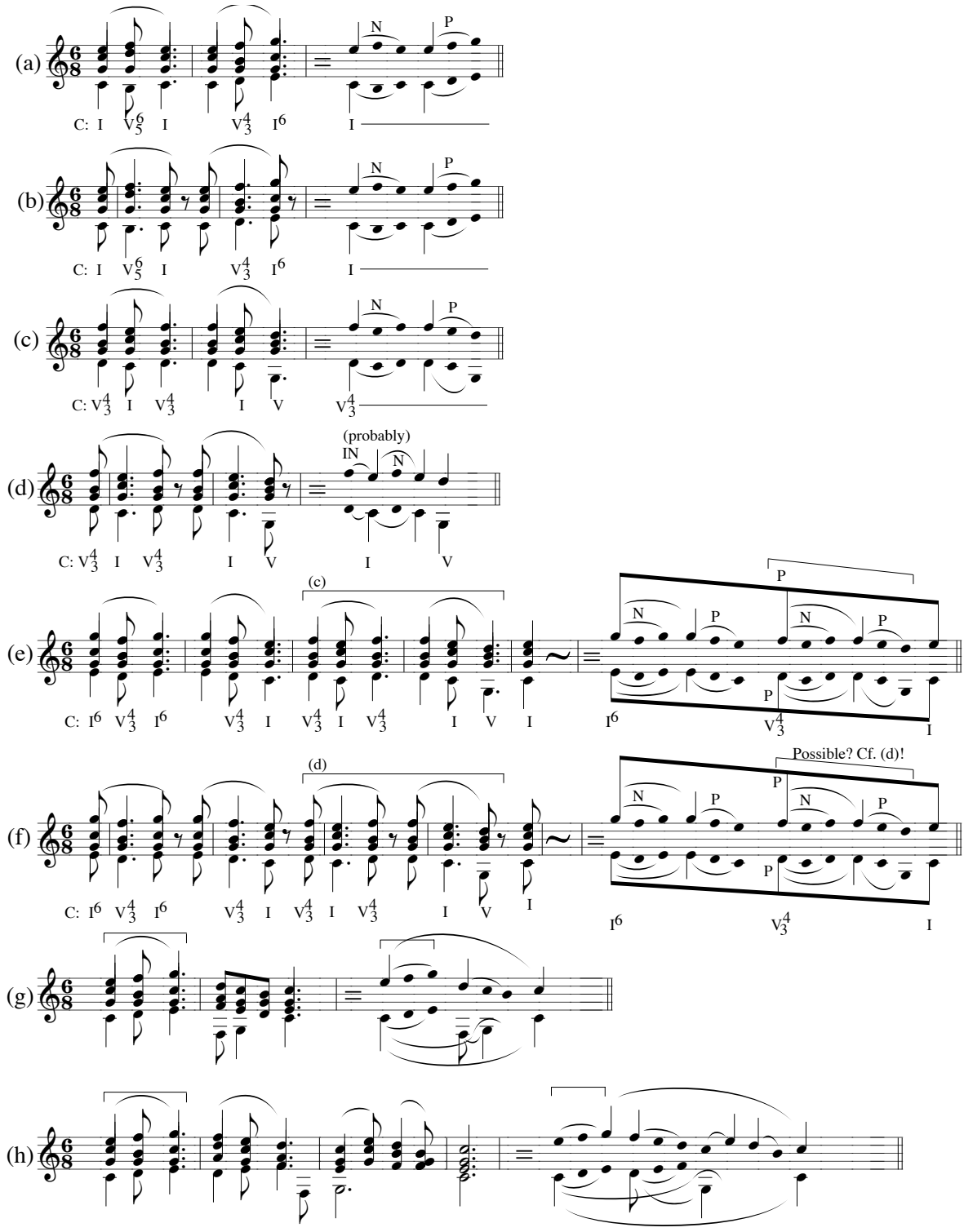
---

[65] Lerdahl and Jackendoff also lucidly demonstrate the significance of both harmonic stability and melodic/rhythmic circumstances for prolongation (much before Larson's and Straus's discussion). See especially Lerdahl and Jackendoff 1983: 118 ff. (this discussion concerns time-span reduction but also pertains to prolongation owing to the interaction between these aspects of organization). Another early discussion pertinent to the present considerations is in Rothgeb 1978, whose Example 3 demonstrates the same thing as the present Example 8c–d.

Moreover, certain successions of melodic intervals offer stronger embellishment clues than others. Let us first consider stepwise progressions. Especially strong clues are given by stepwise three-note successions. Such successions give clues to the existence of passing or neighboring figures, depending, of course, on whether the two steps go to the same direction or back-and-forth. Also larger unidirectional stepwise progressions offer strong embellishment clues. Stepwise two-note successions, on the other hand, give weaker embellishment clues. These observations may be formalized in terms of the schemata consisting of S (a structural element) and E (an elaborative element or a chain thereof) introduced in section 3.1.2, adding a lower-case s to symbolize an element that is structural in relation to E but less structural than S. Hence SE and ES arpeggiations become SEs and sES, respectively, by elaborating them by passing motion. If it is possible to interpret stepwise motion as SES (neighboring motion), SEs or sES (passing motion), strong embellishment clues emerge. However, if only SE and ES are possible—that is, if there are only two elements separated by a step—embellishment clues are weaker, since the melodic motion does not produce a strong bias in favor of one or the other alternative (however, the bias may be strengthened by organizational principles specific to styles or pieces).

As a consequence of these considerations of meter and melodic figures, the functional clarity of incomplete neighbors, accented neighbors, and accented and passing tones depends on harmonic stability more crucially than that of complete neighbors and passing tones at the normative weak metrical position. Example 8a–d offers some illustration. All these fragments comprise two groups of three chords with stepwise top voice, offering embellishment clues to the structural superiority of the framing element. In (a), these clues are supported by rhythm and meter as well as harmonic stability. In (b), they are supported by harmonic stability clues but contradicted by rhythm and meter, which emphasize the *middle* element in each group. In (c), in turn, they are supported by rhythm and meter but contradicted by harmonic stability. Finally, in (d) they are contradicted by rhythm and meter as well as harmonic stability. It would seem that only the last combination is sufficient to outweigh the clues given by grouping and melodic figures, elevating the middle elements in a group to a higher structural status. This clearly demonstrates that both embellishment clues (related to condition #3) and harmonic stability clues (#1 and #2) are significant for prolongation. If the former are strong enough, as in (a) and (c), they may override the latter, but if they are weak or ambiguous, as in (b) and (d)—owing to the conflict between grouping, on the one hand, and rhythm and meter, on the other—harmonic stability becomes decisive (*pace* Larson).

EXAMPLE 8. Harmonic stability and embellishment clues as structural determinants



Returning to different interval successions, combinations of leaps also offer embellishment clues if it is possible to interpret them as arpeggiation. Such clues depend less on the temporal order of elements: both SE and ES occur frequently (whether or not they are

filled in by passing motion to form SEs or sES), with some general bias perhaps in favor of SE. A significant factor in embellishment clues for arpeggiations is register. Tones in registral extremes—higher in the upper voice, lower in the bass—are more likely to assume structural priority; this, of course, is inherent in the notion of structural *outer* voices. For an example in which embellishment clues clarify an arpeggiation, overriding harmonic stability clues, consider the not infrequent instances of ascending arpeggiations towards the primary tone of the *Urlinie* in which the latter is accompanied by an unstable (non-tonic) harmony (as in Example 5o above). Notwithstanding the local instability, the goal tone of the arpeggiation assumes structural significance, partly because the clarity of the arpeggiation justifies interpreting it as connected with the opening tonic.

Different factors of embellishment clues often conflict with each other. For example, in Example 8a, rhythm and meter work in favor of E5, but register in favor of G5. It would seem that the former factors are strong enough to dominate in this case, but similar conflicts often produce ambiguous situations whose interpretation requires considerations of the larger context. That is to say, one has to examine which of the candidates for structurally superior tones participates more essentially in large-scale motions (cf. Larson 1997: 118–9). This is illustrated by the two passages in Example 8g–h. The bracketed figure at their beginning reproduces m. 2 of Example 8a; removing m. 1 weakens the rhythmic and metric emphasis on E5. In this case, the clues favoring E5 and G5 are approximately in balance, and only the continuation tips the balance in either direction. This demonstrates that in determining the structural order within one level we have to allow for embellishment clues formed not only by the tones at that level but also at the next higher level. Also, consider again Example 5o: the structural significance of the goal tone of an arpeggiation towards the primary tone must, according to Schenkerian theory, be confirmed by a subsequent *Urlinie*.

Another very important factor of embellishment clues involves various kinds of parallelisms and motivic relationships. Generally speaking, repetitions of a motive offer a clue that similar structural relationships prevail within different occurrences (although, once again, unequivocal harmonic relationships may override this clue). Examples 8e and 8f implant Examples 8c and 8d, respectively, into a context in which they are heard as parallelistic in relation to the preceding measures. In Example 8e, this further corroborates the interpretation of (c). In Example 8f it seems possible that the parallelism may even alter the original interpretation of (d).

Finally, it should be added that embellishment clues do not only depend on the internal melodic relationships within a voice but also on the contrapuntal relationships between different voices. In conventional tonality, motions in parallel tenths and sixths as well as voice-exchange motions are favored patterns that yield additional embellishment clues. (For example, in Example 5s, the implicit D5 is supported by the parallel-tenth motion it creates in relation to the bass.)

The above considerations of embellishment clues in conventional tonality are largely applicable to the analyses in **I–III** as well. Rhythm, meter, articulation, grouping, different successions of "steps" and "leaps", register, relationships between structural levels, and motivic relationships have largely similar significance in both tonal music and in the present post-tonal examples. As for contrapuntal relationships, the parallel tenths and sixths are often replaced by other parallel intervals, such as the minor sevenths in the introductory section of *Vers la flamme* (Example 7c above). One way to explain such common features would be to appeal to the influence of conventional tonality on early post-tonal music. However, an alternative (or supplementary) explanation is that such factors are to some extent manifestations of general perceptual tendencies utilized in different styles. For example, it seems probable that a "stepwise" figure from an accented A to unaccented B and back will, largely regardless of context, tend to be experienced as governed by A rather than by B (exemplifying an SES schema). The closer details of embellishment clues, on the other hand, vary according to style. Consider, for example, a stepwise descending two-note figure in which the first tone is accented. In some historical periods of conventional tonality (classicism, especially), *appoggiaturas* become so commonplace that such figures might be understood as offering an embellishment clue supporting the structural superiority of the second tone, in other words, an ES schema. The repertoire in **I–III**, on the other hand, includes less appoggiaturas, and a similar figure would more likely be SE.

While the analytical interpretations emerge from the combination of harmonic stability clues and embellishment clues, for reasons of space such factors are not always made explicit in **I–III**. For an example of explicit discussion, one may consult the treatment of mm. 9–12 in Berg's op. 2/2 in **II**, in which the obscurity of harmony is balanced by the clarity of embellishment clues that unambiguously support the present analysis.

3.2.3.4 CONDITIONS #1 AND #4: RELATIONSHIPS BETWEEN HARMONIC AND MELODIC NORMS

Both conditions #1 and #4 involve a classification of intervals. Harmonic intervals are classified as consonances and dissonances (condition #1) and melodic intervals as arpeggiations and voice-leading intervals (condition #4). Insofar as we assume that these classifications derive from the presence or absence of intervals in the referential harmony, we may expect them to be closely related: consonant intervals and arpeggiating intervals are those included in the referential harmony; dissonant and voice-leading intervals are absent from it.

Actually, however, the relationship between these classifications is more complex. Most importantly, the latter distinction is not entirely derivable from the referential harmony but observes, in both tonal music and in all the present examples, some form of the proximity

principle of voice leading: small intervals, "steps," function as voice-leading intervals and larger ones, "leaps," as arpeggiations. In tonal music, for example, the diminished fifth is absent from the triad, which makes it dissonant but *not* a voice-leading interval.[66] As melodic intervals, diminished fifths typically function as arpeggiations within dissonant harmonies such as V^7 (dissonant chords may also be arpeggiated).

The dissonance of an interval is thus not a sufficient condition for its ability to fulfill a voice-leading function but is it a necessary condition? In **II** (section 1), I suggest that this need not be the case. Even if the referential harmony contains small intervals, it may be possible that melodic occurrences of such intervals function as voice-leading intervals, simply by virtue of a syntactical rule which rules out the arpeggiation of such intervals. These considerations suggest an alternative view of the relationships between conditions #1 and #4. Against the above assumption that these classifications both derive from the referential harmony,[67] we might alternatively suggest that they have quite different sources: the construction of the referential harmony for condition #1 and the proximity principle for condition #4.

Considerations of actual music, both conventional tonality and the present post-tonal examples, suggest that the issue is best illuminated by a combination of these alternative views. The referential harmony and the proximity principle of voice leading are autonomous but reciprocally interacting sources of organizational principles. Due to such interaction, the classifications required by conditions #1 and #4 do tend to correlate to some extent. On the one hand, the construction of the referential harmony modifies the application of the proximity principle of voice leading. The tendency of intervals to be heard as voice-leading connectives or arpeggiation depends both on their width and on the possibility to associate them with the referential harmony. On the other hand, the proximity principle of voice leading may influence the choice of referential harmonies. If the referential harmony excludes "stepwise" intervals, an especially clear distinction between arpeggiations and voice-leading intervals becomes possible because "stepwise" intervals cannot be associated with the intervals in the referential harmony. Insofar as such clear distinction is aesthetically desirable, the exclusion of "stepwise" intervals enhances the utility of a referential harmony.

Let us first discuss the ways in which the referential harmony may modify the application of the proximity principle of voice leading. Two aspects are significant for this issue: the borderline between small and large intervals, and the extent of permissible octave generalizations. While we may identify a "default" borderline, going between 2 and 3 semitones (for its perceptual justification see section 5.3), it is open to modifications that depend partly on the referential harmony. For example, if there is a whole-tone in the referential harmony, this produces a greater tendency for melodic whole-tones to be associated with that in

---

[66] There are other aspects of complexity not considered here. For example, the consonance or dissonance of an interval does not depend solely on the interval but also on its position within a harmony. For example, in conventional tonality, the fourth is dissonant against the bass but consonant between upper voices.

[67] This assumption is also evident in Straus's (1987) discussion of analytical examples.

the harmony. Sometimes this association is powerful enough to justify their interpretation as arpeggiation. Debussy's *Ce qu'a vu le vent d'ouest* is a case in point. In **III**, I have interpreted this prelude on the basis of an "added sixth" referential harmony F♯–A♯–C♯–D♯ and its semitonal variants, which all contain a whole-tone, either C♯–D♯ or C♮–D♮. Since these chords are transposed to other levels only in small-scale prolongations, these specific whole-tones occupy a special position in the prelude. Consequently, the association of melodic whole-tones with harmony is the most powerful for these two whole-tones. This offers justification for interpreting whole-tones according to their transpositional level: C♯–D♯ and C♮–D♮ function as arpeggiations but other whole-tones as voice-leading intervals, an assumption confirmed by its descriptive power for both the large-scale events and details.

The influence of the referential harmony on the extent of octave generalizations permitted by the proximity principle of voice leading was already touched upon in section 3.1.1. In short, if the referential harmony excludes all instances of interval classes 1 and 2, as in conventional tonality, larger realizations of these interval classes ("sevenths" and "ninths") more readily substitute for "stepwise" voice-leading intervals. If the referential harmony excludes simple semitones and whole-tones but includes "sevenths" and "ninths," as in several of the present examples, such substitution is less viable since melodic "sevenths" and "ninths" may also stand for arpeggiations. However, these principles do not always apply rigidly. In triadic tonality, sevenths by no means always stand for voice-leading intervals but often arpeggiate seventh chords. In fact, even a second may function as an arpeggiation within a seventh chord (e.g., the C4–D4 in Mozart's K. 545, m. 15; see Example 3 above). This demonstrates that particular contexts may decisively influence the interpretation of an interval with respect to the harmony/voice-leading condition. Conversely, in some of the present post-tonal examples, a "seventh" or "ninth"—the latter more readily (see considerations of *ro-intervals* in section 4.1.2)—may substitute for a "step" under a clarifying context.

Now let us move to consider the ways in which the proximity principle of voice leading may influence the choice of the referential harmony. The issue of octave generalizations pertains to this issue as well. The productivity of the triad as a referential harmony for prolongational structures may partly be explained by the clarity of the harmony/voice-leading distinction enabled by the exclusion of steps between *pitch classes*. This, in effect, is what Straus (1987: 5) suggests in the following: "voice leading in tonal music proceeds from one pitch-class to another pitch-class adjacent within the diatonic collection (that is, one step away). Harmonic intervals are formed by non-adjacent elements within the collection. From this point of view, the special place of the triad can be clearly understood—it is the maximal subset of the diatonic collection consisting entirely of non-adjacent elements."

In applying similar considerations to the present post-tonal examples, it is generally necessary to substitute pitches for pitch classes, in other words, to abandon octave generalization. Moreover, "adjacency" within a referential collection is substituted by

"proximity," the small size of an interval. Thus, the exclusion of adjacent pitch classes is replaced by the simple exclusion of semitones and whole-tones in the spacing of referential harmonies, a principle called hereafter the *proximity principle of spacing*. This principle holds in most of the examples in **I**–**III**, but there are partial exceptions to it (as evident from the above considerations of *Ce qu'a vu le vent d'ouest*). In addition, some of the present examples show some kind of partial octave generalization in their avoidance of "stepwise" intervals in the referential harmony. In *Voiles*, the opening (mm. 1–20) is based on a harmony (Q) consisting of the pedal B♭ and the "augmented triad" A♭–C–E (Example 4: chord Q). While the pitch-class relationships *between* these two elements contain "steps" (A♭–B♭ and B♭–C), the internal relationships within the "augmented triad" do not. Moreover, since the opening of *Voiles* is exclusively based on the whole-tone set, we may readopt the notion of "adjacency" and explain the status of the "augmented triad" on exactly same basis as the (major or minor) triad in diatonic tonality: it is the maximal subset of the whole-tone collection consisting entirely of non-adjacent elements.

Schoenberg's op. 19/2 manifests another kind of partial octave generalization. A central principle in this piece involves the use of "major sevenths" (registrally ordered interval 11) in consonant harmonies and the exclusion of the "minor ninths" (larger realizations of registrally ordered interval 1). Since the latter is more closely associated with the semitone, the strongest voice-leading interval, this principle, evident also in some other pieces by Schoenberg, may manifest the need to buttress the harmony/voice-leading distinction (**I**: 237).

Above, it has been suggested that the avoidance of "steps" or small intervals or their octave-extensions in referential harmonies may be motivated by the resulting effects on the clarity of the harmony/voice-leading distinction. It should be added, however, that an alternative or supplementary explanation can be based directly on the properties of small intervals as simultaneities. Both horizontal and vertical "steps" have psychoacoustical special properties. Whereas the proximity principle of voice leading relates with the psychoacoustical *streaming* phenomenon, the proximity principle of spacing relates with effects created when two tones are within a critical band. Hence the avoidance of small intervals in a referential harmony adds to psychoacoustical support for *both* the consonance–dissonance and the harmony/voice-leading distinction, and it is difficult to tell which is more important. This issue will be taken up in sections 5.1 and 5.3.

3.2.3.5. CONDITION #2 IN RELATION TO #3 AND #4: "SCALE DEGREES" AS BY-PRODUCTS OF CONTRAPUNTAL EMBELLISHMENT?

Above, the four conditions have been divided into two pairs: conditions #1 and #2 concern harmonic stability, whereas #3 and #4 concern melodic relationships. However, one may ask

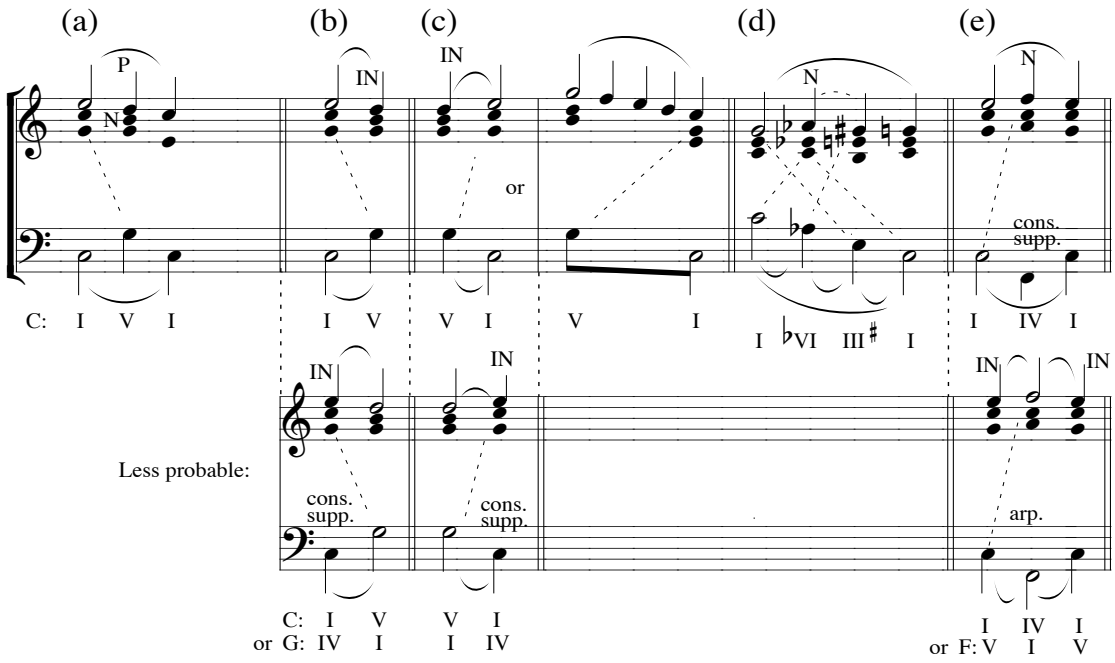
whether condition #2 is not, in fact, a corollary of melodic (i.e., horizontal) relationships in harmonic progressions, as regulated by conditions #3 and #4. In conventional tonality, the scale-degree hierarchy is based on the I–V–I motion in the bass, which arpeggiates the fifth of the tonic triad. This is characteristically elaborated by intervening II, III, or IV (or II^6), harmonies whose bass notes Schenker (1935/1979: § 53 ff.) calls "space-fillings." In terms of the above-discussed categories of embellishment (section 3.1.1.), II may be identified as anticipatory arpeggiation within V; III as further arpeggiation within I; and IV as an incomplete neighbor of V. In upper voices, the bass arpeggiation is characteristically counterpointed by stepwise upper-voice motions: passing $\hat{3}$–$\hat{2}$–$\hat{1}$, i.e., the *Urlinie*, and neighboring $\hat{1}$–$\hat{7}$–$\hat{1}$, as in Example 9a below (cf. Example 5g).

All in all, the scale-degree hierarchy may thus be seen as evolving through the contrapuntal embellishment of the tonic chord, in which the bass arpeggiation is the most decisive element. In order to apply similar considerations to the present post-tonal examples, it is useful to identify two simple factors that pertain to the superordinate position of the I in the normative I–V–I progression of tonal music. The first factor is based on pitch and may be called the *inclusion factor*. The inclusion factor favors I over V since the pitch classes in the bass line I–V–I are included in I but not in V. This means that the bass line may be understood as embellishing I, but not V, by arpeggiation. The second factor is based on temporal relationships and will be called simply the *temporal factor*. It favors I in the I–V–I progression because I occurs at both ends of the progression. In terms of the schemata introduced in section 3.1.2, the temporal factor favors SES over ESE, in accordance with the above discussion on "embellishment clues."

While in the normative I–V–I progression the structural superiority of I is established by the cooperation of the two factors, the individual impact of each of them may be illuminated by considering cases in which one of them is "neutralized" or made equivocal, or in which the two factors contradict each other. Example 9b–e depicts such cases, showing inclusion relationships by dotted lines.

The temporal factor is "neutralized" by removing the tonic from either temporal end, producing SE (I–V) and ES (V–I) schemata (Example 9b–c). Such "incomplete" progressions are significant and occur frequently at all structural levels except for the very highest. At the highest level, i.e., as frameworks of entire pieces of movements, they occur only exceptionally: I–V in some Baroque preludes (whose status as "entire pieces" is questionable) and V–I in some Romantic compositions based on the auxiliary cadence. In any case, the significance of the inclusion factor is evident in the fact that these progressions are much more significant than those in which the structural order is the reverse, as indicated in the lower system.

EXAMPLE 9. The impacts of the inclusion factor and the temporal factor on scale-degree hierarchy



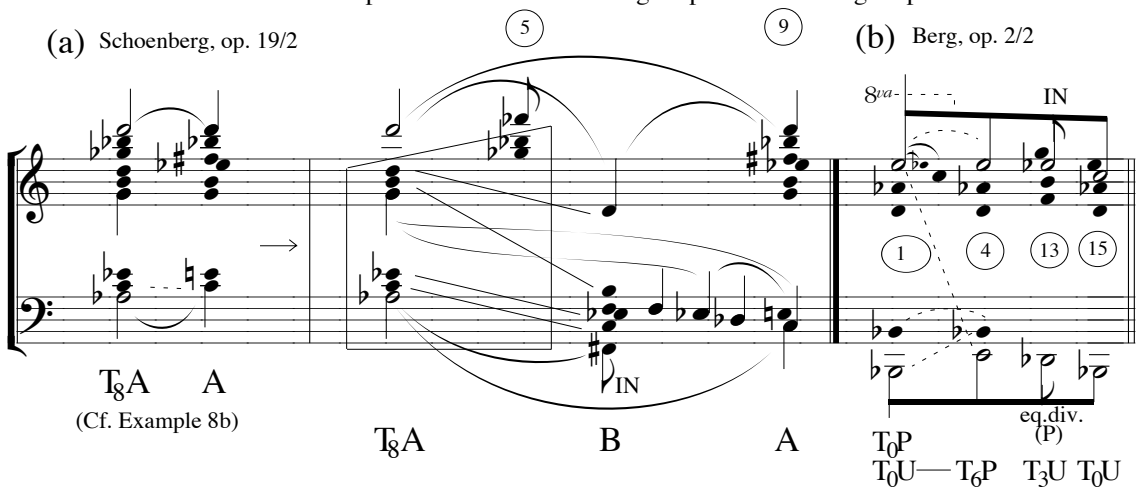
For an example in which the inclusion factor is "neutralized," one may consider progressions based on equal divisions of the octave (Example 9d). In such a case, the temporal factor becomes decisive: there is no plausible alternative for considering the framing harmony as superordinate.

Finally, in Example 9e, the two factors conflict with each other. The inclusion factor favors the middle element, but the temporal factor favors the framing element. In such a case, the temporal factor tends to dominate, even though this means that the bass line does not form embellishment within the governing harmony (cf. the discussion of Example 5l–m above). The "ESE" schema, or more precisely a combination of ES and SE, as shown in the lower system, would require exceptional emphasis on the middle element. However, the significance of the inclusion factor is corroborated by the fact that the I–IV–I progression, which conflicts with the inclusion factor, is structurally much less important than I–V–I; it does not serve as a background structure.

The "scale degree" systems in the present studies may be determined on the basis of the same factors. Interestingly, however, none of them shows the two factors cooperating in the unequivocal manner of the normative tonal progression I–V–I; hence one or the other becomes decisive. This would seem to reflect the tendency of "post-tonal" aesthetics to avoid overly emphasized hierarchical relationships, a tendency related to the discussion of centricity in section 3.1.3 above. Example 10 shows two examples. In Schoenberg's op. 19/2 (Example 10a; cf. **I**: Examples 14–16), the inclusion factor supports the superiority of the opening harmony on

A♭ (chord T8A) over the concluding one on C (chord A); the resulting schema is SE. The former harmony includes the bass of the latter, C. Other events in the bass and in the lowest inner voice (G–E♭–C) are also readily relatable to the opening harmony. On the other hand, there are several features in the music—such as registral manipulation—that impart a stabilizing, "toniclike" quality to the *concluding* harmony. This may seem to contradict the present interpretation, but as I suggest in **I** (248–51), these two views may be reconciled by recognizing the difference between prolongation and centricity.[68]

EXAMPLE 10. Inclusion and temporal factor in Schoenberg's 0p. 19/2 and Berg's op. 2/2



In Berg's op. 2/2 (Example 10b; cf. **II**: Example 11c), the inclusion factor is "neutral" in a way foreign to conventional tonality. This stems from the fact that the referential harmony (T_0U) includes a bass-related tritone, which is horizontalized in the background bass arpeggiation (B♭–E–B♭). Since transposing a tritone by a tritone keeps pitch classes invariant, the harmony on E (T_6P)[69] also includes the pitch classes of the bass arpeggiation. However, reading T_0U and B♭ as superordinate to T_6P and E is supported by the temporal factor (and also by the lower registral position of B♭).[70]

It may be interesting to compare the Berg example with the introductory section of Scriabin's *Vers la flamme*; see Example 7c above. These examples have a common referential harmony, chord U, and give special emphasis to the subset of U indicated as P; in conventional terms, chord P corresponds to the "French sixth," whereas U enlarges P by adding a major ninth. In both cases, the bass line is based on a tritone arpeggiation and includes an equal

[68] The identification of the transpositional levels in symbols T8A and A reflects the course of discussion in **I**, which starts from the assumption that the concluding harmony is referential, even though it ultimately turns out to be prolongationally subordinate to the opening one.

[69] Since P is a subset of U, the relationship between T_0U and T_6P is based on transposition in accordance with the present definition of "scale degrees" (see note 62).

[70] The Berg example is complicated by the fact that the T_0U at the end of the structure is not the end of the song. The conclusion of the song is forms a transition to op. 2/3; see **II**: Example 11.

division of the tritone as an elaborative element. (As observed above in section 3.2.3.1, the U chord is only locally referential in *Vers la flamme*. Strictly speaking, the introductory section of *Vers la flamme* does not embody an SES schema but SE, T_0U–T_6U, which is followed by the establishment of another locally referential harmony. However, as explained in **II**, this new harmony is a superset of T_0U in terms of pitch classes and shares the root E with it, and may therefore be perceived as substituting for the return of T_0U; see also Example 22a,v below. The structural superiority of the T_0U and of the bass E in the bass arpeggiation of the introductory section is also supported by the concurrent upper-voice passing motion, D4–C4–B♭3. In the overall organization, the superiority of the E bass is conclusively confirmed by its powerful return in m. 95.)

While the roles played by the inclusion factor and the temporal factor vary, they are sufficient to determine the structural order of "scale degrees" in these examples; more or less similar considerations apply to other examples of the present studies. If the transpositions of the referential harmony produces motions—usually in the bass[71]—that may be interpreted as arpeggiation within that harmony, considerations of the "scale degree" hierarchy are based on such arpeggiation.[72]

These considerations give reason to return to the question whether the scale-degree condition can be treated as a corollary of the embellishment condition and the harmony/voice-leading condition.[73] As a kind of counterargument, one might remark that from the experiential or phenomenal standpoint the notion of "scale degrees" as harmonies of different stability based on transpositional levels is not quite reducible to harmonies brought about by the contrapuntal embellishment of the governing harmony. An important practical consequence that this distinction implies for musical structures is that "scale degree" systems might be based on transpositional schemes that do *not* link with the embellishment of the referential harmony. Hence even if condition #2 follows from conditions #3 and #4, the opposite relationship does not hold: there may be music that fulfills the former condition but not the latter conditions, with respect to bass lines in chord progressions. In the present studies this issue is raised in

---

[71] Exceptions are given by some episodes in Debussy's *Ce qu'a vu le vent d'ouest*, which contain small-scale "scale degree" systems based on *upper-voice* arpeggiation (**III**).

[72] The upper-voice motions that counterpoint the bass arpeggiation vary from case to case: in the Scriabin passage (Example 7c), the D4–C4–B♭3 passing motion may be compared to an *Urlinie*; in the Schoenberg and Berg examples (Example 11) the bass arpeggiation does not support similar passing motion.

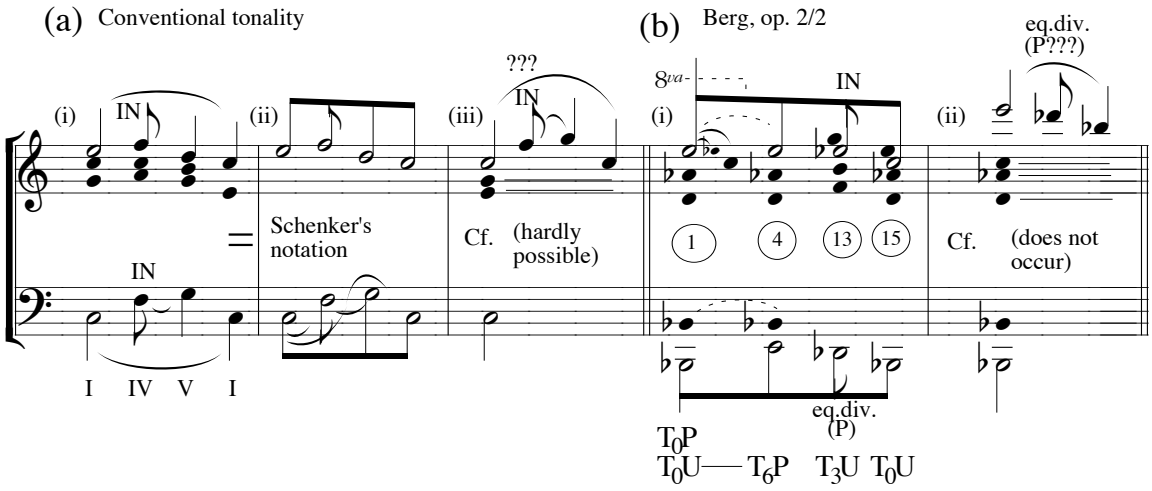
[73] One possible argument for a more independent status of "scale degrees" might be that while the establishment of a "scale degree" system requires interpretations based on conditions #3 and #4, *after* such establishment, "scale degrees" independently affect structural relationships. For example, the establishment of the tonic level in conventional tonality requires motions based on the embellishment of the tonic (such as I–V–I, V–I, or I–V), but after such establishment the tonic bears superior stability that has autonomous impact as a clue to structural relationships. However, the notion of such autonomous impact is somewhat questionable. Tonic (or "tonic") triads frequently occur as non-structural chords (i.e., in nontonic functions); they seem to be no more resistant to fulfilling passing or neighboring functions than any other chords, if the local factors related to conditions #3 and #4 support such functions.

connection with equal divisions of the octave, a technique employed in parts of *Vers la flamme*. I shall return briefly to this issue at the end of this subsection.

There are, however, practical reasons to pay special attention to condition #2 even when the "scale degree" system is coordinated with motions embellishing the governing harmony—as is normally the case in both conventional tonality and in the present post-tonal examples. This is because the embellishment figures in the bass lines of chord progressions tend to differ somewhat from "ordinary" embellishments, i.e., those occurring at the surface without chord change. This means that even if we view condition #2 in terms of #3 and #4, the application of the latter to the bass lines of chord progressions has to be treated as a special case.

Consider the most commonplace bass line in conventional tonality, I–IV–V–I. The bass of the IV was above identified as an incomplete neighbor (see Example 11a,i); however, as "ordinary" embellishment (as in Example 11a,iii) such an incomplete-neighbor figure would be most unusual. The special nature of such bass motions is also reflected in Schenker's discussion (1935/1979: §53 ff.). He did not identify the pre-dominant bass notes in terms of "ordinary" embellishment categories (as was done above) but discussed them as special types of "space-fillings." As is well known, he also employed special notational symbols for such motions (Example 11a,ii).

EXAMPLE 11. Differences between bass-line embellishments and "ordinary" embellishments



Bass-lines tend to deviate from "ordinary" embellishments also in the examples of the present studies. Consider the equal division of the tritone in the Berg example (Example 10b, 11b,i; similar considerations also apply to the *Vers la flamme* passage in Example 7c.)[74] As

[74] Schoenberg's op. 19/2 (Example 10a) also includes a bass-line feature uncharacteristic of surface embellishments, namely, the substitution of "major ninth" A♭3–F♯2 for a whole-tone as a voice-leading interval. Contextual support for this relationship is given by the clarity of the registral transfers between several upper voices (C, E♭, B, D) in chords on A♭ and F♯ (chords A and B). However, according to the present definition (note

discussed in **II**, it is possible to interpret the E2–D♭2–B♭1 progression as passing motion, by modifying the borderline between small and large intervals so that pitch-interval 3 ("minor third") becomes a voice-leading interval. The plausibility of such an interpretation is enhanced by the absence of interval class 3 from the referential harmony. Nevertheless, similar voice-leading 3s do not occur at the surface level in the way shown in Example 11b,ii. As in the conventional I–II–V and I–IV–V motions, the basic arpeggiation is not prolonged by an element of "ordinary" embellishment but by a special kind of "space-filling" technique.

These considerations suggest that bass lines of chord progressions tend to involve liberties uncharacteristic of "ordinary" embellishments, especially concerning the use of larger intervals in non-arpeggiating function. For explaining such a tendency, three points seem worth making. First, bass lines of chord progressions may involve such liberties because they are less liable to create confusion with respect to the harmony/voice-leading distinction. Consider the hypothetical Example 11b,ii, in which the equal division of the tritone occurs as foreground figuration. It seems ambiguous whether the figure stands for passing motion or arpeggiation, in other words, whether the D♭ belongs to the harmony or not. In the bass line of the actual Berg song (Example 11b,i), a similar equal division causes no ambiguity, since foreground occurrences of the referential harmony and its transpositions (chord U) make it amply evident that D♭ (fb-interval 3 or "minor third" in relation to the governing B♭) is non-harmonic.

Second, having discussed what makes bass-line liberties possible, some considerations are warranted as to what makes them useful. For understanding this issue, it should be observed that harmonic progressions serve several concurrent purposes, only one of which is the linear embellishment of the governing harmony in the bass. Chord choices may be seen as brought about through the accommodation of such different purposes; liberties in bass-line embellishments facilitate such accommodation. One purpose already discussed is supporting upper-voice motions; Examples 5l–m and 9d–e show bass lines fulfilling this function without constituting any embellishment themselves. In I–II–V and I–IV–V progressions, II and IV provide the opportunity to give consonant support to the $\hat{4}$, a function especially significant in pieces based on the $\hat{5}$-*Urlinie*. In Berg's op. 2/2, the T_3U, which yields an equal division of the tritone in the bass, offers support for the top-voice motion E–E♭–C, an enlargement of the opening foreground motive (Example 11b,i). However, the scope of this explanation is limited. For example, in $\hat{3}$-*Urlinie* progressions, the IV and II are unable to provide any additional support for *Urlinie* tones, but are nonetheless about as common as in $\hat{5}$-*Urlinie* progressions. A noteworthy additional consideration is variation of pitch-class content. The II and IV are the two triads that include the two pitch classes in the diatonic scale—$\hat{4}$ and $\hat{6}$—that lie outside I or V. In the Berg example, the two primary harmonies, T_0U and T_6P are very close to each other in terms of pitch classes (the latter is a subset of the former), but T_3U produces variety; it also

62), chord B is not a "scale degree," since it is not obtained by transposition but involves altering chord construction.

brings about a change in the otherwise prevalent transposition of the whole-tone set. Hence we also may count the complementation of pitch-class resources among the factors relevant to chord choices.

Third, the tendency to use larger intervals in bass lines—a tendency dating back to Renaissance music—may have some kind of a psychoacoustical correlate in the fact that the width of the critical band becomes larger in lower registers. Accordingly, in a certain psychological sense, "leaps" appear to be "smaller" in the bass (**I**: 244 [n. 53]).[75] This could, in part, explain the non-arpeggiating use of large intervals in progressions such as those in Example 11 (and, also, in Example 5l–m). (However, even in the bass, such progressions do not occur as foreground figuration without supporting chord progressions.)

If bass-line embellishments tend, for all these reasons, to deviate from "ordinary" embellishments, one may ask whether such a tendency could be strong enough to completely annul the connection between the two. Progressions through equal divisions of the octave may be understood in this light. Such equal divisions occur in both conventional tonality (see, e.g., Schenker 1935/1979: Fig. 100,6) and in more modern music; in the present studies they are discussed in connection with *Vers la flamme* (**II**). Unless the governing harmony contains similar equal divisions—as in some music of Liszt based on the augmented triad (Cinnamon 1986)—such equal divisions cannot be interpreted as embellishment of the harmony. Sometimes such progressions may be explained as giving support to stepwise upper-voice motions (Example 9d) but such an explanation is not always equally viable. We may regard equal-division progressions as being borderline cases of prolongation. Owing to the temporal factor, there is no ambiguity as regards the governing harmony, but the voice leading does not always meet conditions #3 and #4 in the ordinary sense. One way to view such techniques is to define them as a special way to embellish octaves in chord progressions, which cancels the normal harmonic implications of horizontal intervals.[76] However, owing to the dichotomy between the governing harmony and the horizontal framework (cf. Cinnamon 1986), such progressions represent a "weaker" form of prolongation than those in which the latter "composes out" the former.

---

[75] For a general discussion on the psychoacoustical significance of the critical band and the somewhat divergent results concerning its width in different registers, see, for example, Hartmann 1997: 249–58.

[76] Such a definition is related with Lester's (1971) notion of *division tone*, a non-harmonic tone dividing harmonic intervals (not only octaves) in equal parts without (necessarily) relating stepwise with harmonic tones. According to Lester (ibid.: 6), "division tones are particularly useful at middleground levels" but "are not frequently used at the immediate foreground, where they would add to the number of hanging pitches." Hence Lester observes a characteristic difference between middleground and foreground embellishments that corresponds to the present differentiation between bass-line embellishments and "ordinary" embellishments.

3.2.4 SUMMARY

The above discussion of conditions for prolongation may be summarized as follows. There are two basic requirements for prolongation. The first is a consistent system of *harmonic* stability, based on chord construction (condition #1) and transposition (condition #2). It seems justified to "liberalize" this requirement by permitting stability systems in which these two factors play roles that deviate from conventional tonality (as described by Straus). The second basic requirement is a consistent system of *melodic* relationships, which involves a well-defined set of embellishment figures between tones of different weight (condition #3) and the distinction between arpeggiations and voice-leading intervals (condition #4). Structural relationships emerge from the combination of clues given by harmonic stability and melodic relationships. If a relatively unstable harmony is prolonged at a certain level, it must be subordinate to a more stable harmony at a higher level. The realms of harmonic and melodic requirements overlap with respect to chord progressions, which involve both harmonic stability (condition #2) and melodic relationships (#3 and #4).

4. CONCEPTION OF HARMONIES AND INTERVALS

Whereas the present concept of prolongation is, with the specifications discussed above, based on Straus's conditions, the ways of applying these conditions to post-tonal music deviate from those of Straus 1987. There are two main differences. First, the conception of harmonies and intervals is different. Whereas Straus's considerations are based on standard set-theoretical concepts, I argue for the significance of aspects ignored by such concepts: registration and rootedness. The second difference concerns the factors taken into account in considerations of norms relevant to the conditions (such as consonance and dissonance). As illustrated in Example 1 above, the present approach allows for both psychoacoustical factors and the general context, whereas Straus is only concerned with the latter factor.[77] These two issues relate with each other: one of the main arguments for allowing for aspects outside standard set theory is the perceptual, psychoacoustical significance of such aspects. Despite their relationships, I shall treat these two issues separately. Section 4 argues for the significance of registration and rootedness, considering perceptual issues on an informal basis, and discusses relevant theoretical concepts. Section 5 then discusses more specific connections between psychoacoustic phenomena and functional norms.

4.1. REGISTRATION

4.1.1. THE CASE FOR RESTRICTING OCTAVE EQUIVALENCE

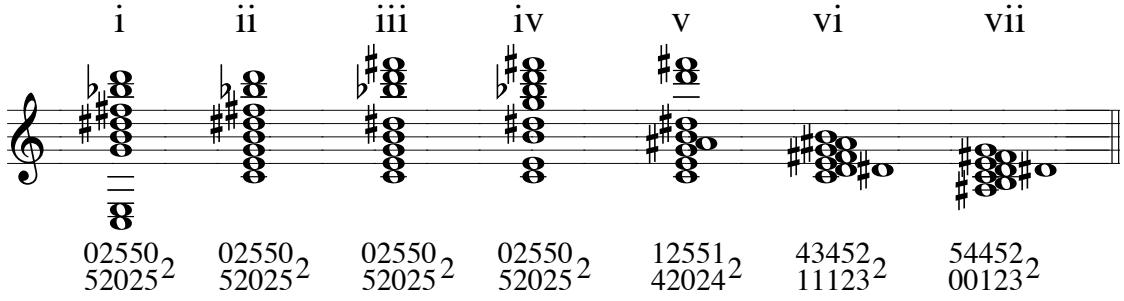
The relevance of registral distinctions for the conditions of prolongation may be justified by several arguments, including both *a priori* and *a posteriori* reasoning. The latter is based on our knowledge of specific cases of prolongational structures. In both conventional tonal structures and in the examples of the present studies, functional norms depend, to variable extents, on registral distinctions. Allowing for such distinctions is crucial for revealing prolongational structures in both cases. In terms of the model in Example 1, the "general context" in each case supports functional distinctions based on registral distinctions: the general treatment of harmonies depends on their registral features. In conventional tonality, this clearly applies to

[77] Straus does not discuss this issue in depth, but the confinement to contextual considerations is evident in his discussion of analytical examples.

5/3, 6/3, and 6/4 chords; for a demonstration concerning post-tonal music, see especially **I**: 231–3.

There are also *a priori* reasons to assume that registration pertains to functional norms in post-tonal as well as tonal prolongation. Such *a priori* reasons have both logical and perceptual grounds. The larger the referential harmony, the less there tend to be logical possibilities for basing the functional distinctions required by conditions #1 and #4 on interval classes. Straus (ibid.: 8) observes this problem in his discussion of Roy Travis's (1966) analysis of Schoenberg's op. 19/2: "Travis's sonic sonority, set-class 8–19 (01245689), contains every interval class at least twice. As a result, it is impossible to interpret the voice leading motions." ("Sonic sonority" is a misprint; it should be "tonic sonority," a notion that corresponds to the present "referential harmony.") In general, the more pitch classes there are in a referential harmony, the more probably it contains instances of all six interval classes, which therefore cannot offer a clear basis for functional distinctions. However, from this we cannot infer that large harmonies are unable to offer the basis for functional distinctions, but merely that such distinctions cannot observe unrestricted octave equivalence. Treating intervals in a register-sensitive manner, for example, by replacing the six interval classes by the eleven *registrally ordered intervals*, greatly increases the logical possibilities to form large referential harmonies with an interval content manifesting such distinctions. For example, while the Schoenberg sonority (the concluding harmony of op. 19/2; see Example 12i) includes all six interval classes, it excludes registrally ordered intervals 1, 5, and 9. (See **I**: Example 2. See also the vector under the present Example 12i. The notation in this 11-entry vector stems from Castrén 1997. The numbers of ro-intervals 1–5 are set in an upper row from left to right, the number of ro-interval 6 appears on its own on the right, and the numbers of ro-intervals 7–11 are placed in a lower row from left to right; hence complementary intervals are aligned vertically.)

EXAMPLE 12. Registral rearrangements of the concluding harmony of Schoenberg's op. 19/2



The exclusion of registrally ordered interval 1 (semitone or "minor ninth") from the Schoenberg chord has particularly significant ramifications for the organization in op. 19/2. More generally, registral distinctions tend to be especially significant for interval classes 1 and

2, owing to the special status of semitones and whole-tones as evident in the proximity principle. The significance of registral distinctions for the logical possibilities of constructing referential harmonies becomes particularly clear by considering the options for referential harmonies that exclude these intervals (cf. section 3.2.3.4). In a registrally insensitive conception of harmony, such exclusion means that interval classes 1 and 2 do not occur at all in the referential harmony, which leaves room only for five such options with cardinality greater than 2: the four traditional triads types and the diminished seventh chord (set classes 3-10, 3-11A, 3-11B, 3-12B, and 4-28 in the transpositional classification). Needless to say, making distinctions based on registral ordering or pitch-intervals greatly increases such options.

As for perceptual arguments that support the significance of register for functional norms, I will not discuss at length here those that derive from the specific ways in which psychoacoustical phenomena correlate with functional norms; these ways will be discussed in section 5. In passing, it may be noted that the relevant psychoacoustical phenomena depend on registration; hence it is easy to understand why such dependance is also more or less evident in the corresponding functional norms. What makes this argument less conclusive, however, is that such dependance tends to some extent to be diminished in functional systems due to cognitive octave generalizations. While *a posteriori* considerations of existent functional systems show that such generalizations do not amount to complete freedom of registration, it is hard to see *a priori* grounds for excluding such a possibility.

A more conclusive and generally applicable perceptual argument draws simply on perceptual recognizibility. Whether or not the functional norms show any correlation with psychoacoustical properties in the specific ways to be discussed in section 5, they should be based on perceptual aspects that can be recognized in a sufficiently immediate way. Consider the basic assumption that functional consonance is based on the similarity with the referential harmony. Theoretically, one may identify different aspects of similarity, some of which have a more direct relationship with the immediately perceptible properties of harmonies than others. And, as is one of the basic arguments in **I**, the register-insensitive aspect of similarity based on pitch-class sets is often not recognizable in a sufficiently immediate way, whereas register-sensitive non-set-theoretical aspects often are.

To avoid misunderstanding, it should be noted that I do not wish to suggest that all theoretical or analytical concepts should describe "immediately perceptible properties" in an equally direct way. (cf. section 2.3 above). The present studies, for example, point out numerous "concealed" features whose relationship with immediate perception is more indirect and complex (but which are, I believe, quite significant for the experience of qualities such as "coherence," or "organicism"). However, concepts with a relatively simple relationship with immediate perception and concepts with a more complex relationship play a different role in musical organization. An important function of prolongational syntax is to support and clarify "concealed" relationships, but such support and clarification presupposes that functional

norms regulating prolongational syntax—such as consonance and dissonance—are themselves based on more immediately perceptible factors. Even if the following considerations suggest that such factors cannot usually be based on set-theoretical concepts of similarity, they do not rule out other kind of analytical utility of such concepts.[78]

The argument of perceptual recognizibility may be presented in a weaker or stronger form. The weaker form states that distinctions based on registration are recognizable in a sufficiently immediate way and therefore *may* pertain to syntactic norms. For example, we can immediately tell a fifth from a fourth, or a 5/3 chord from a 6/4 chord, which shows that the distinction between inversionally related intervals may pertain to consonance and dissonance. (This is, of course, confirmed, *a posteriori*, by considerations on conventional tonality.) The stronger form states that registral distinctions *must* pertain to functional distinctions since the similarity between different registral realizations of a given pitch-class set is often not immediately recognizable.

Whereas the weaker argument is generally applicable, the stronger argument depends on the pitch-class set in question, especially on its cardinality. For example, the stronger argument does not seem applicable to triadic tonality: both the similarity and the difference between 5/3 and 6/4 chords would appear to permit immediate recognition. Hence *this* argument cannot rule out a functional system in which these chords bear similar function, even though such a system is not realized in conventional tonality. For larger pitch-class sets, however, the similarity between different registrations tends to become increasingly difficult to recognize. Example 12 (a reproduction of **I**: Example 7) demonstrates this for the concluding eight-note harmony of Schoenberg's op. 19/2. Chord i shows the actual registration, whereas chords ii–vii are different registrations of the same pitch-class set. On the basis of informal evaluation, the similarity between chords i–iv seems immediately recognizable, whereas that between, say, chords i and vi does not. Of course, we can speak of "immediate recognizibility" only in a relative sense; moreover such a property also depends on the capacities of the listener. Nevertheless, it seems safe to maintain that the recognition of the similarity between chords i and vi in Example 12 is considerably more difficult than, say, that between any registrations of a triad. More generally, the larger the harmony, the more difficult it tends to be to recognize its similarity with other registral realizations of the set class it represents.[79] To explain this tendency, suffice it to note here that large harmonies, in general, pose greater difficulties for perception than small ones; hence there is a greater need to restrict such difficulties by lesser registral variation.

Example 13 (a reproduction of **I**: Example 8) presents another kind of comparison. In terms of pitch-class sets, chords A and B (8-19 and 6-Z13) are not closely related, but they share features of intervallic make-up. These features include approximate registral density and,

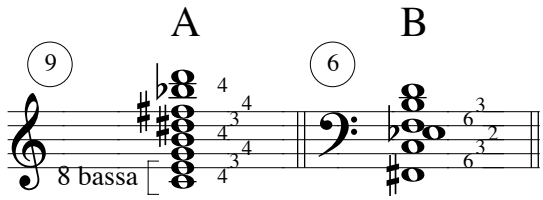
---

[78] Cf. the discussion of prolongational and motivic-associational aspects of organization in **I**: 233-34.

[79] This tendency should be understood only in a rough sense. Even a relatively large pitch-class set may be fairly easily recognizable in different registrations if it has a very distinctive interval-class content; consider, e.g., the whole-tone set (6-35).

most significantly, the consistent use of registrally ordered interval 11 (“major seventh”) between registral layers. This is indicated by the numbers beside the chords indicating intervals between registally adjacent tones; each chord is based on a cyclic generator, 4+3+4 or 6+3+2, with the total width of 11. On the basis of informal evaluation, it would seem that, in comparison with the set-theoretical similarity between chords such as i and vii in Example 12, the common features between chords A and B in Example 13 are fairly evident for immediate perception. (For a more formal treatment of chord construction in A and B, see **I**: 241, rules (i)–(iii); for a survey of chords based on these rules, see ibid.: Example 9.) This suggests that the latter aspect of similarity is more likely to function as the basis of functional properties such as consonance and dissonance, a suggestion confirmed, *a posteriori*, by the analysis of Schoenberg's op. 19/2. Thus, set-theoretical similarity is neither a sufficient nor a necessary condition for perceptual or functional similarity. It should be added, however, that while chords A and B both function as consonances in the Schoenberg example, chord B is a consonance of a lesser rank than A, in accordance with the status of A as the referential harmony (cf. the considerations on gradated consonance in section 3.2.3.1 above).

EXAMPLE 13. A comparison between two chords in Schoenberg's op. 19/2



	A	B
ro-interval vector	02550 52025 2	02202 30220 2
ic vector	545752	324222
FB intervals	2,3,4,6,7,10,11	5,6,8,9,11

The arguments based on both logical possibilities and perceptual recognizibility suggest that the significance of registral distinctions for functional norms tend to increase rather than decrease for music with larger referential harmonies. Because registral distinctions also pertain to functional norms in triadic tonality and because post-tonal music tends to employ larger harmonies than triads, this gives general reasons to assume that the register-insensitive approach of standard set theory is inadequate for considerations of prolongation in post-tonal music.

4.1.2. SOME THEORETICAL CONCEPTS

If prolongational considerations thus require register-sensitive concepts of harmonies and intervals, what kind of concepts should be adopted? A point of departure for considering this issue is given by Robert Morris's (1995) three candidates for pitch-set equivalences: PSC, PCINT, and FB (**I**: 238 ff.). Under PSC (for "pitch-set class"), chords are equivalent only if they are identical or exact transpositions of each other, with no registral rearrangements. Under PCINT (for "pitch-class INT") they are equivalent if the successions of registrally ordered intervals between registrally consecutive tones—pitch-class INTs—are identical. Pitch-class INTs are shown in Example 13 by the numbers to the right of chords A and B. PCINT may also be defined as a requirement that the *registrally ordered pitch-class sets* are identical or transpositionally related. It permits octave enlargements and reductions of intervals but not changes in registral ordering. Under FB (for "figured bass"), chords are equivalent if they have identical sets of bass-related registrally ordered intervals. FB permits free octave equivalence among the upper voices but not changing the bass (i.e., chord inversion in the traditional sense). In Example 12, no chords are equivalent under PSC. Chords i–ii are equivalent under PCINT. Chords i–vi are equivalent under FB.

PSC, PCINT, FB, and, finally, pitch-class-set equivalence under T_n present a continuum of possibilities in which the freedom of registration gradually increases. (This continuum could be supplemented by other kinds of partial orderings besides FB.)[80] These equivalences involve different interval concepts. Under PSC, all *pitch-intervals* remain invariant. Under PCINT, all *registrally ordered intervals*—abbreviated *ro-intervals* in the present studies—remain invariant. Under FB, all bass-related ro-intervals, to be called *fb-intervals*, remain invariant. All these interval concepts may be denoted on the basis of semitones. Pitch-intervals are denoted by the amount of semitones separating the two pitches. Ro-intervals are obtained from pitch-intervals by subtracting octaves (12s) off; hence there are, apart from 0 (octave/unison), eleven ro-intervals (for a formal definition, see **I**: 232, note 8). (The 11-entry ro-interval vector has already been introduced in connection with Example 12.) The six interval classes, in turn, may be obtained from ro-intervals by equating complementary intervals, that is, by substituting 12–i for i when $i > 6$. In the present studies, such semitone-based notation is used in **I** and **II**, whereas **III** uses traditional interval names.[81]

The concepts of PCINT and FB reflect the notion that all intervals are not equally significant within a pitch set or chord. PCINT involves a special focus on the intervals between

[80] For one example, consider a "polychord" harmony consisting of registral layers, in which the mutual order of the layers remains invariant, but the internal order of tones within a layer varies (**I**: 239).

[81] Enharmonic distinctions are left outside consideration throughout the present studies. The traditional significance of such distinctions stems from the functional system of conventional tonality and is therefore less relevant for the present considerations. Admittedly, enharmonic orthography may link with structural features also in less traditional circumstances (see, for example, Perle 1984). However, insofar as the validity of structural interpretations is measured according to experiential relevance (for the listener), enharmonic spelling cannot determine structural relationships but at most confirm the interpretation based on audible factors.

registrally consecutive tones, FB on those formed with the bass. Both perceptual intuition and well-established musical practices support the special significance of such intervals. The special status of the bass (as in FB) is, of course, familiar from conventional tonality and also evident in the present examples. From the perceptual point of view, such special status is connected with the special status of the fundamental in the harmonic series. The task of relating a group of pitches to a low pitch is familiar to audition from the perception of complex tones and their virtual pitches. Such a relationship between chord perception and complex-tone perception is especially significant for the property of rootedness, an issue to be taken up in sections 4.2 and 5.2. Trends of 20th-century music often involve a shift from FB towards PCINT. The special significance of intervals between registrally adjacent tones is evident, for example, in the concept of "quartal harmonies."[82]

Besides bass-related intervals and adjacent-tone intervals, other kinds of intervallic relationships may also enjoy special status in some circumstances. As an example, consider the consistent use of ro-interval 11 between registral layers in chords A and B of Example 13.

The ways in which octave equivalence should be regulated vary largely in different kinds of music. In the present studies, the extent of registral constraints lies, generally speaking, somewhere between FB and PCINT. These concepts are not always applied in an absolutely rigid way; some informal flexibility has been allowed. In the analysis of Schoenberg's op. 19/2 (**I**), PCINT offers the point of departure for the conception of harmony, but limited registral reorderings are allowed without change of functional status.[83] In **II**, which focuses on harmonies related to the harmonic series and on the property of rootedness, FB is of primary significance, reflecting the governing position of the lowest tone in the harmonic series. However, whereas in **I** registral constraints are somewhat informally relaxed from PCINT, in **II** they are somewhat informally tightened from FB: Some registrations of upper voices are much more characteristic than others. A special position is occupied by harmonies which reproduce the registral ordering in the harmonic series (that is, relate to the series by PCINT) or even the actual pitch-intervals (PSC) (cf. sections 5.2.1–2 below). In **III**, the conception of harmony is treated in a more informal manner. The analysis of Debussy's *Ce qu'a vu le vent d'ouest* identifies two aspects of organization, one based on semitonally related harmonies and another based on the outer-voice structure (see Example 23b below). As for the former aspect, no formal restrictions are given for the registration of harmonies (although important tendencies exist), but the outer-voice structure is interpreted in relation to the lowest bass note, in

[82] A famous occurrence is in the introduction of Schoenberg's Chamber Symphony no. 1; also see the theoretical description in Schoenberg 1911/1922/1978: 399 ff. Concepts based on adjacent-tone intervals are also discussed in such a practically oriented book as Persichetti 1962.

[83] For example, the similarity between chords i–iv in Example 12 is evident enough to serve as the basis of similar function. Incidentally, their similarity is reflected in their identical ro-interval vectors, but I would not lay too much importance on such a feature. The descriptive powers of ro-interval vectors is diminished by the fact that they only depict the overall ro-interval content, ignoring differences in the position and significance of different intervals.

accordance with FB. Finally, it should be added that, owing to the impact of both aspects of the proximity principle (see section 3.2.3.4 above), registral distinctions tend, in general, to be especially significant for the functional status of interval classes 1 and 2.

4. 2. ROOT SUPPORTS

The psychoacoustical underpinnings of rootedness and root supports will be discussed in section 5.2. At present, it suffices to give a theoretical definition of the concept of *root supports*, on the basis of the lowest range of the harmonic series. The present definition follows Parncutt 1988 in all but one respect. In the present definition, the range relevant to roots comprises the first eleven harmonics, going one further than in Parncutt's model (for discussion on this topic, see **II**: section 1.1).

A root support may be defined as an interval that approximates, in equal temperament, an interval between the fundamental and a harmonic within the relevant range (the first eleven harmonics in the present studies). Such a definition may be applied to either pitch-intervals or ro-intervals, but in the present studies (following Parncutt 1988) root supports are normally understood as ro-intervals. Moreover, considerations of root supports are usually confined to bass-related ro-intervals, i.e., fb-intervals. (As an exception to this one may mention the discussion of the total ro-interval content of the concluding harmony of Schoenberg's 19/2 [**I**: 237–8]; the discussion of "secondary root supports" in section 5.2.4 below also involves the possible significance of root-supporting intervals between upper voices.)

The root-supporting *weight* of an interval depends on how low in the harmonic series the approximated harmonic occurs: the lower the harmonic, the greater the weight of the corresponding interval. These definitions yield the following list of root-supporting ro-intervals (fb-intervals), from the strongest to the weakest (excluding 0, i.e., octave/unison): 7, 4, 10, 2, 6, or, in conventional terms, fifth, major third/tenth, minor seventh, major second/ninth, tritone; for illustration, see Example 15b below.

The concept of root supports defines a dimension of interval properties which, unlike set-theoretical interval concepts, is not based on the width of intervals measured in semitones. The significance of this dimension is supported by arguments partially similar to that of registration (subsection 4.1.1). The *a posteriori* argument of analytical productivity is supported by several examples in **II** and **III** which show a clear correlation between functional and expressive properties, on the one hand, and the property of being root-supporting, on the other. The argument of perceptual recognizibility is also applicable, but only in the weaker form. Harmonies consisting of root supports (see **II**: Example 1), share certain recognizable characteristics, which enhances the comprehensibility of functional systems based on such harmonies. However, regarding post-tonal prolongation in general, there may be cases in which

the property of being root-supporting plays no consistent role; hence this aspect has less universal significance than registration.

Finally, it may be observed that the considerations on root supports also bear on the preceding issue, the restriction of octave equivalence. In particular, the root-supporting weights imply a strong contrast between the strongest root supports, fb-intervals 7 (fifth) and 4 (major third) and their inversions 5 (fourth) and 8 (minor sixth). This contrast is evident in the stability system of tonality, especially with respect to 7 and 5 (a perfect consonance) and the fourth (a dissonance). I shall return to this and other issues relevant to rootedness in section 5.2.

5. CONNECTIONS BETWEEN PSYCHOACOUSTICS AND MUSICAL ORGANIZATION

I shall discuss the correspondence between three psychoacoustical phenomena and three principles of musical organization. The psychoacoustical phenomena are critical band, virtual pitch, and streaming; each of them represents a major aspect of auditory perception.[84] The respective organizational principles are the proximity principle of spacing, rootedness, and the proximity principle of voice leading. The former two pertain to condition #1 (consonance–dissonance). The latter pertains primarily to condition #4 (harmony/voice-leading), but owing to the relationships between the conditions (see sections 3.2.3.2 and 3.2.3.5) it also has ramifications for conditions #2 and #3.

I shall illuminate each of these correspondences by describing the psychoacoustical phenomenon, the principle of musical organization, its applicability in the present repertoire and, for the sake of comparison, in conventional tonality. In each case it should be borne in mind that, despite the correspondences, the psychoacoustical phenomenon and the music-theoretical principle cannot be equated. The partial correspondence with psychoacoustical phenomena helps to establish comprehensible principles of organization, but, as discussed in section 2, such principles also depend on the general context and on cognitive generalizations and extensions.

5.1 CRITICAL BAND, ROUGHNESS, AND THE PROXIMITY PRINCIPLE OF SPACING

Critical band is a concept with multifarious ramifications for pitch perception.[85] If two pure-tone frequencies are closer to each other than the critical band, their interference produces effects of *roughness* and *masking*, the former of which is often identified as "sensory dissonance." From the middle register upwards, the width of the critical band corresponds roughly to pitch-interval 3 (minor third); in lower registers it is wider.

The proximity principle of spacing (already introduced in section 3.2.3.4) is the principle

---

[84] The significance of these psychoacoustical principles for post-tonal prolongation is also discussed in Lerdahl 1999. Lerdahl's approach deviates from the present one in that according to Lerdahl "for atonal music the chief measure of tension is psychoacoustic at all levels," (ibid.: 23). In terms of the model in Example 1, this means that Lerdahl pays less attention to the impact of the "theoretical" level as mediating between psychoacoustics and analysis.

[85] For a general discussion of the significance of the critical band and the somewhat divergent results concerning its width in different registers, see, for example, Hartmann 1997: 249–58.

of avoiding small pitch-intervals, 1 and 2 (semitone and whole-tone), in consonant harmonies. This principle has an obvious connection with the critical band. For two tones violating the principle, critical-band effects emerge between all harmonics of the same order number, including, of course, the fundamentals.

The proximity principle of spacing has more or less systematic significance in most of the present examples. It does not amount to the main principle of consonance in any of them but may often be interpreted as a supplementary principle (see, for example, **I**: 241, Rule (i); **II**: section 1.3).[86]

Considerations on critical-band effects, as related to the proximity principle of spacing, also shed some light on the argument of perceptual recognizibility, as discussed in reference to Example 12 above. For example, owing to the cluster of semitones and whole-tones, chord vi produces quite different (much more drastic) critical-band effects than chord i.

The proximity principle of spacing also prevails in conventional tonality, since consonant harmonies (5/3 and 6/3 chords) include no seconds. However, as already observed, an important distinctive feature between conventional tonality and the present post-tonal examples is that the former kind of organization observes full octave generalization of this principle: triadic consonances exclude not only simple semitones and whole-tones but all occurrences of interval classes 1 and 2. Such octave generalization is at least partially cognitive rather than directly based on psychoacoustics. To be sure, simple sevenths and ninths tend to produce a relatively high amount of critical-band effects, arising between the higher tone's harmonics of order number n and the lower tone's harmonics of order number 2n. However, greater registral separation makes considerations of critical-band effects irrelevant but does not cancel the functional dissonance of interval classes 1 and 2.[87]

None of the analytical examples in the present studies manifests such full octave generalization of the proximity principle of spacing. However, the consonance–dissonance relationship between ro-intervals 11 and 1, as evident in some music by Schoenberg (**I**), might stem from a PCINT type of partial generalization (section 4.1.2). That is to say, the consonance of ro-interval 11 ("major seventh") in relation to 1, even when the latter are registrally enlarged ("minor ninths"), may be supported by the cognitive association of such 1s with the simple semitone, the interval which produces maximal critical-band effects. It may be noted, however, that a similar consonance–dissonance relationship between ro-intervals 10 and 2 is not evident in any of the present examples. To explain this difference, it may suffice to note that cognitive extensions of psychoacoustical phenomena require that the original phenomena are rather

[86] A notable counterexample is Debussy's *Ce qu'a vu le vent d'ouest*, in which the whole-tone C♯–D♯ between upper voices of the referential harmony is established as a consonance through contextual reinforcement (**III**).

[87] For example, if the distance between two tones is about four octaves, the upper tone lies in a register in which the lower tone's harmonics are situated at intervals of about one semitone; moreover, they are usually inaudible. Hence there is no difference between the critical-band effects induced, say, by pitch-intervals 48 (four octaves) and 49 (four octaves + semitone).

strong, and that critical-band effects are considerably weaker for whole-tones than for semitones.[88]

Regarding all these considerations of the correspondence between critical-band effects and the proximity principle of spacing, it should be added that the avoidance of small pitch-intervals and their octave-relatives in harmonies may also partly stem from the desire to enhance the harmony/voice-leading distinction (cf. section 3.2.3.4). This distinction is based on the proximity principle of voice leading, which is psychoacoustically related with the phenomenon of stream segregation (section 5.3). The principle of avoiding small intervals in harmonies has thus two important perceptual consequences, the relative importance of which is not easy to determine. For verticalities, the principle helps to diminish critical-band effects. For arpeggiated harmonic intervals, it helps to avoid their integration into one stream, which would conflict with their arpeggiating function.

It should also be added that the proximity principle of spacing (with possible octave generalizations) by no means accounts for all musical ramifications of critical-band effects. In calculating the total amount of roughness produced by two tones one must allow for all pairs of harmonics. With tones of normal harmonic spectra, roughness tends to be minimized by the simplicity of frequency ratios—slight mistunings of such ratios, as in the equal-tempered fifth, increase roughness only slightly and do not cancel such a tendency.[89] For example, a fifth (pitch-interval 7, which approximates frequency ratio 3:2) characteristically induces less roughness than a tritone (pitch-interval 6), although neither interval violates the proximity principle of spacing. Such an order, of course, concords with the functional status of these intervals in tonal music. However, since such differences correlate less consistently with functional stability in the present post-tonal repertoire, they will be left outside discussion here.

5.2 VIRTUAL PITCH AND ROOTS

5.2.1 PSYCHOACOUSTICAL BASIS AND THEORETICAL CONSIDERATIONS

Virtual pitch is a pitch percept synthesized from the group of harmonics that form a complex tone. In the synthetic mode of hearing (the normal way of hearing musical tones), we do not hear the harmonics separately but the virtual pitch, which corresponds to the fundamental frequency.[90] The auditory system has a highly developed capacity to retrieve virtual pitch, on

---

[88] See the consonance curve in Plomp and Levelt 1965: Fig. 10.

[89] Ibid.: Fig. 11.

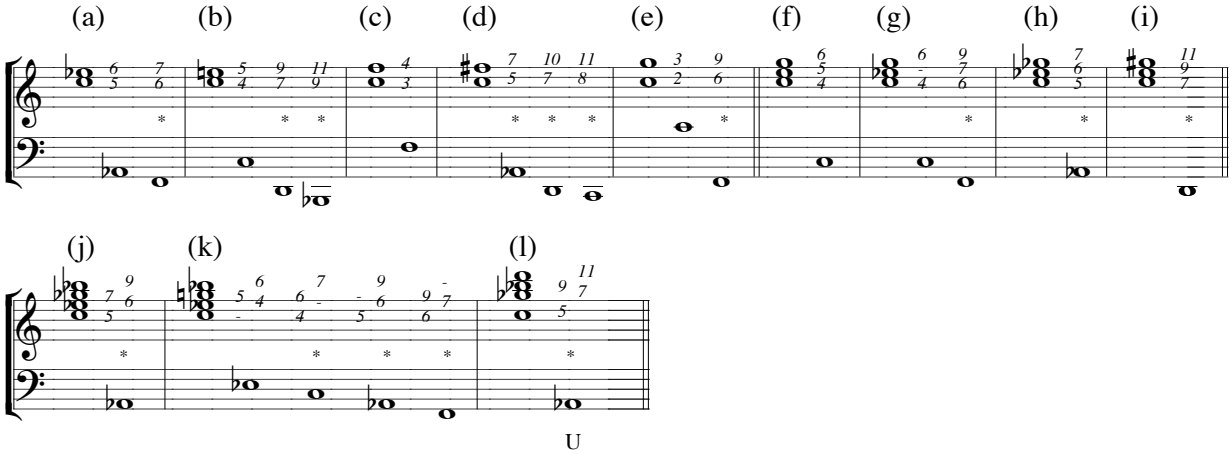
[90] Besides "virtual pitch," other names, e.g., "residue pitch" are used for the same phenomenon in psycho-acoustical literature. I use "virtual pitch" following Terhardt, who made the original suggestion concerning the significance of this phenomenon for musical harmony. Besides Terhardt's theory of virtual pitch, based on pattern recognition, there are several other explanations of the phenomenon. Moreover, there is experimental evidence for the existence of multiple mechanisms for retrieving virtual pitch (see, e.g., Houtsma and

the basis of just a few harmonics, even when these do not include the fundamental frequency.[91]

As suggested by Ernst Terhardt (1974, 1982), this capacity is also utilized in the perception of chords. To the extent that the intervals between the bass and the upper tones in a chord are similar to those between the fundamental and partials in a harmonic series, the bass will have a tendency to be heard as a root, that is, governing the overall pattern in a manner of a virtual pitch. Since virtual-pitch perception permits considerable mistunings of harmonics—up to at least a quarter-tone (Moore et al. 1985; cf. Parncutt 1988, 70–71)—its relevance to rootedness is not canceled by the use of equal temperament.

To get an intuitive idea of the perceptual significance of roots, one may consider Example 14. Listen to the pitch sets (intervals and chords) on the upper stave and experiment with combining them with different bass tones along the chromatic scale. The lower stave shows bass notes that the virtual-pitch theory predicts to have a greater or lesser tendency to be perceived as roots. In other words, the upper-stave pitch sets fit (wholly or partially) to the approximated harmonic series of the indicated basses, as shown by italicized numbers. The degree of rootedness depends on root-supporting weights, that is, on the closeness of the approximated harmonics. Rootedness does not necessarily presuppose that all upper voices correspond to harmonics; such non-correspondence is shown by a "minus" sign (-) in Examples 14g and 14k. (Except for these cases, Example 14 concentrates on bass notes permitting the total fit of upper voices to the harmonic series, omitting other root candidates.)

EXAMPLE 14. Root candidates for sets of upper voices



Smurzynski 1990, Winkler et al. 1997). Ultimately, the question of what kind of mechanisms pertain to virtual-pitch perception is less crucial for the musical consequences.

[91] For example, Houtsma and Goldstein (1972) showed that virtual pitch may be retrieved on the basis of two harmonics presented separately to each ear.

For the present purposes, the most important feature demonstrated by Example 14 is the possibility to create non-traditional (= non-triadic) root effects by using correspondents of harmonics higher than the first five (or their octave relatives); such possibilities are indicated by asterisks. These effects rely on weaker root supports than the traditionally consonant triads. Both the major and the minor triad include the strongest non-octave-equivalent root support, the fifth (fb-interval 7); the major triad also includes the second strongest, the major third (fb-interval 4). In the examples of the present studies (**II**–**III**) the weakness of root supports tends to be compensated for by a greater number of them. For example, Example 14(l) shows the U chord familiar from previous Debussy, Scriabin, and Berg examples (Examples 4, 7c, 10b). This chord excludes fb-interval 7 (the fifth) but includes all other root supports, that is, approximations of the first eleven harmonics. (The restriction to eleven harmonics corresponds approximately with the range most relevant to virtual pitch.)[92]

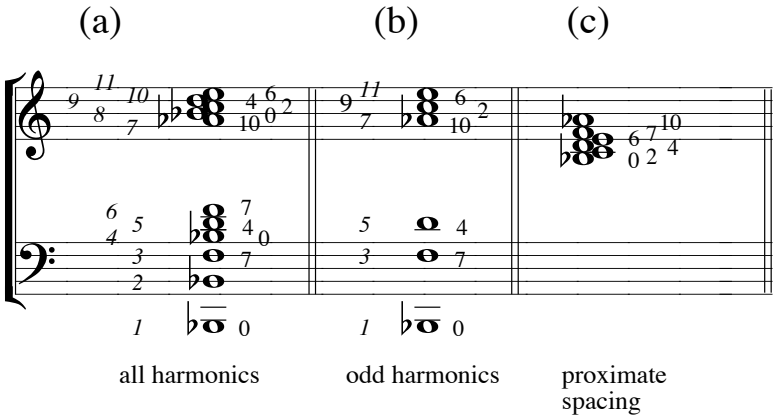
Example 14 is based on pitches with specific registral location and thus ignores the kind of octave generalization that is manifest in the conventional conception of roots and in the concept of root support as defined in section 4.2. In this conventional conception, roots are pitch classes rather than pitches. For example, the C major triad is understood as being in root position as long as its lowest tone is C, regardless of register. Moreover, there are no restrictions concerning the registral relationships between the upper voices. In terms of Morris's equivalence types, the conventional concept of roots is based on an FB conception of harmony (section 4.1.2).

While the FB conception also offers the point of departure for considerations of rootedness in the present studies, it is not entirely sufficient for the treatment of the issue. In comparison with conventional tonality, the registral ordering of upper voices, and even pitch-intervals, become more important for the identity and rootedness of harmonies (**II**: 6 ff.). This reflects partly the general tendency, observed in section 4.1.1, that owing to the largeness of post-tonal harmonies their recognizibility requires stricter registral constraints. Consider the chords in Example 15. The italicized numbers to the left of the chords indicate correspondences with the harmonics of the bass; the ordinary numbers to the right indicate fb-intervals. Example 15a comprises approximations of the first eleven harmonics of B♭1. Example 15b comprises approximations of odd harmonics only, and thus excludes octave repetitions. Example 15c, in turn, realizes the fb-intervals of the first two chords within one octave. While all three chords are equivalent in terms of FB, the task of recognizing Example 15c as similar to 15a or 15b is considerably more difficult than the task of recognizing the similarity between any registrations of a root-position triad (cf. the discussion in reference to Example 12). For larger harmonies,

[92] For example, Houtsma and Smurzynski (1990. 309) divide harmonics into "those of low order and those of high order, with the separation somewhere between the 10th and 13th harmonic," on the basis of the more distinct virtual pitch produced by the former.

there is less perceptual justification for positing full octave equivalence among upper voices and more reason to move in the direction of PCINT or PSC: the perceptual support for rootedness is strengthened by reproducing the registral ordering or even the actual pitch-intervals of the harmonic series. As regards registral ordering, it should be observed that the present examples (in **II**) often favor the odd-harmonic registration (Example 15b). The favored position of this spacing is partly explained by the fact that it helps to avoid violations of the proximity principle of spacing between correspondents of harmonics 7–11 (cf. Example 15a).

EXAMPLE 15. Different registrations of root supports



The significance of registration for rootedness is also increased by the use of weaker root supports: it would seem that such weakness needs to be balanced by stronger registral faithfulness. A strong root support such as fb-interval 7 (the fifth) is so strongly established in our audition and cognition that it permits more extensive octave generalization than weaker root supports such as 2 (major ninth) and 6 (tritone). Moreover, it may be observed that the stronger root supports, fb-intervals 7 and 4 (the fifth and the major third) occur in more than one octave in the range of harmonic series that pertains to virtual-pitch perception (Example 15a). The weaker root supports, fb-intervals 10, 2, and 6, only occur once. Hence we are accustomed in virtual-pitch perception to hear 7 and 4 in different registral orderings, whereas the mutual order of 10, 2, and 6 is fixed. This may partly explain a similar tendency in musical organization.

All this suggests that FB is not a totally satisfactory basis for considerations of rootedness. However, the present studies do not attempt to formalize any general requirements for stricter registral constraints, but are limited to making case-by-case observations of PCINT and PSC type relationships (**II**). Positing more general requirements is difficult because such constraints vary according to different musical purposes. Moreover, as discussed in section 2.2.1, cognitive octave extensions that lead far from the harmonic-series spacing may be supported by “concretizations” within individual pieces. In Debussy's *Voiles*, fb-intervals 4

and 10 (D and A♭) occupy the highest registral position in the overall structure (Example 4c), which would seem to contradict with their correspondence with the lowest harmonics of B♭1 within the governing U chord (the odd-harmonic spacing of U is B♭1–D4–A♭4–C5–E5; see Example 4a). However, registrations close to this spacing are found in surface harmonies at strategically important points, such as the first (albeit non-structural) occurrence of U in m. 10–11 (**II**: Ex. 16b). As already discussed, an "organic" musical process connects the original D4 in m. 10—the approximated fifth harmonic of B♭—with the structurally decisive D7 (Example 4e). Regarding the musical purpose of the high registral position of D and A♭, we may observe that it helps to clarify their primary structural significance. This significance in turn reflects their greater root-supporting weights (in comparison to C and E); hence, paradoxically, a property stemming from the lowest position in the harmonic series is reflected by the highest position in the artistic design—but the connection between these opposite positions is concretely manifest in the music.

The present studies also do not attempt to quantify rootedness in a precise way. As explained, rootedness is enhanced by the use of root-supporting fb-intervals, depending on their root-supporting weights, and further enhanced by closer registral connections with the harmonic series (in terms of PCINT and PSC). The present studies do not go beyond these general principles in determining the degree of rootedness. It is beyond the present concerns to try to determine how one should allow for the combined impacts of these factors in order to calculate a numerical measure of rootedness. It should be observed that while Parncutt (1988) discusses verbally the significance of upper-voice registration, his numerical root algorithm is based on an FB-type of conception. Such an algorithm produces results that are fairly well in accordance with the root conception of conventional tonality but is somewhat less adequate for the present considerations. One might attempt to modify Parncutt's root model in a more register-sensitive direction but I do not attempt to tackle here the difficulties that such an attempt involves. It should be observed that rootedness—the extent to which upper tones are understood as belonging to the pattern governed by the bass—also depends on contextual factors; hence the descriptive power of context-free numerical values would remain limited. In any case, such values would not decisively strengthen the present analyses. Insofar as rootedness pertains to the harmonies in the present studies, such pertinence can be demonstrated by rather simple means: such harmonies consist totally or almost totally of root supports (in terms of FB) and are often registrated in a manner more or less similar to the harmonic series (PCINT and PSC).

Finally, the position of non-root-supporting fb-intervals (1, 3, 5, 8, 9, 11) warrants some comments. The inclusion of such intervals in chords by no means negates the significance of rootedness. An important principle of registration is that such intervals have the least "root-detractory" effect on rootedness when they occur above the root supports (Parncutt 1988: 87 ff.). However, these intervals are not entirely equal with respect to their effects on rootedness. Fb-intervals 5 (fourth) and 8 (minor sixth) are inversions of the two strongest root supports. It

would seem that such intervals have a special "root-opposing" effect: they strongly point to the *upper* tone of the interval as the root. As suggested by Parncutt (1996: 72), the conflict between the implied root and the actual bass pertains to the instability of these intervals in conventional tonality, an explanation especially relevant to the dissonance of the fourth.[93] The remaining fb-intervals—1, 3, 9, and 11—are "neutral" in this respect: they represent interval classes (1 and 3) not found between the fundamental and the approximations of the first 11 harmonics.

Another special issue (brought up in **III**) is whether an interval could be understood as indirectly supporting the bass as the root, if it is a root support of a lower tone which, in turn, is a root support of the bass. For example, if a chord contains fb-interval 7 and, above it, 11, the latter is a strong root support (4) of the former, which in turn is a strong root support of the bass. This issue will be elaborated in section 5.2.4.

5.2.2 THE SIGNIFICANCE OF ROOTEDNESS IN THE PRESENT STUDIES

Rootedness has variable significance for the construction of referential harmonies in the present analytical examples. Consequently, there is variable correlation between rootedness and functional consonance.

In the analytical examples of **I**, such correlation is less crucial than in other examples. In Schoenberg's op. 19/2, the referential harmony, chord A in Example 13 above,[94] includes all the root-supporting fb-intervals (2, 4, 6, 7, 10) and, in addition, 3 and 11. Hence the "root-opposing" 5 and 8 are among the excluded fb-intervals (1, 5, 8, 9). Moreover, the registration of chord A has features in common with the harmonic series (see **I**: Example 6). However, functional consonance in this piece (as formalized in **I**: 241) is not based on fb-intervals but on rules concerning intervals within textural layers, on the one hand, and between layers, on the other. Such rules are fulfilled by another chord, chord B in Example 13, which has quite different fb-intervals (5, 6, 8, 9, 11). Nevertheless, some amount of correlation between rootedness and stability is evident in the fact that chord B functions as a consonance of a secondary rank in relation to A.

I (237–8) also observes that the total ro-interval content of chord A is maximally biased in favor of the strongest root supports 7 and 4, as opposed to their inversions 5 and 8. However, the perceptual significance of such a bias is less clear. Some illumination to the issue might be given by a conception based on multiple roots: some of the upper voices function as secondary roots for those above them. In the present studies, ideas of multiple roots are not systematically developed, except for subsection 5.2.4 below, which speculatively discusses "secondary root

[93] Fb-intervals 5 and 8 are also absent from all the referential harmonies in the analytical examples of **I**–**III**.
[94] To be precise, chord A is not the primary transposition of the referential harmony in the prolongational structure of Schoenberg's op. 19/2 (see the discussion in section 3.2.3.5).

supports," defined as root supports of root supports.

II concentrates on examples that show a clear correlation between consonance and rootedness. Referential harmonies consist solely of root supports except for Scriabin's *Vers la flamme*, in which fb-interval 9 is added above them (in accordance with the above-discussed principles of registration). Consonance rules in **II** are based on fb-intervals and closely correspond with the property of being root-supporting. (However, parts of Webern's op. 3/1 manifest supplementary rules based on relationships between registral layers, similar to those in the preceding Schoenberg example.)

III also discusses examples in which rootedness pertains to referential harmonies (including material relevant for the discussion of secondary root supports in subsection 5.2.4 below). In addition, it presents examples in Debussy in which adding a root to a set of upper voices creates special expressive effects without necessarily establishing a referential harmony in the structural sense.

5.2.3 ROOTEDNESS IN CONVENTIONAL TONALITY; SOME REMARKS ON HISTORICAL EVOLUTION

Let us next consider how the two above-discussed psychoacoustical factors of consonance, critical band and virtual pitch, pertain to the consonance system of conventional tonality.[95] One way to explain that system on the basis of these phenomena is the following. First, the intervals that produce the greatest critical-band effects (roughness), minor and major second and the tritone, are excluded, together with their octave generalizations. This leaves the major and minor triad as the only possible three-pitch-class collections within the diatonic set. Second, the registral position of the interval class most strongly determinative of virtual pitch, i. e, 5 (fifth/fourth), determines the stability of different triad inversions. If this interval class occurs as fb-interval 7 (fifth), in accordance with the harmonic series, the chord (5/3) is fully stable. If it occurs as fb-interval 5, whose "root-opposing" quality was discussed above, the chord (6/4) is unstable. If it occurs between upper voices, the stability of the resulting chord (6/3) is between these extremes.

Given that the concept of rootedness has its origins in conventional tonality, its pertinence to the traditional consonance system, as evident in the above explanation, may almost seem surprisingly slight. No root support besides the strongest—the fifth, or fb-interval 7—needs to be invoked in this explanation. However, while the position of interval class 5 is indeed decisive for the basic consonance system of conventional tonality, there are some more subtle indications

[95] Explaining consonance systems on the basis of these two factors follows Terhardt's (1984) "two-component theory of musical consonance." About other purported explanations of consonance and dissonance, see, for example, the discussion in Plomp and Levelt 1965.

that composers were also sensitive to the effects of weaker root supports. Let us first consider fb-interval 4 (major third). Since this interval is included in the major triad but not in the minor, the former is more strongly rooted than the latter. While this difference is not reflected in the basic consonance rules—major and minor triads are both stable—it is in several ways reflected in compositional practices, and also in the expressive connotations of major and minor. As for compositional practices, it may be first noted that altering minor chords to major by mixture is much more significant than the reverse procedure. The dominant occurs regularly as a major chord in minor, and the major tonic often occurs in endings (Picardy third). The reverse phenomena hardly ever occur. Modulations from minor to major keys (usually to the III) are also more significant than the reverse. All in all, as observed by numerous theorists, the major and minor triads are not entirely "equal" in compositional practices; there is a certain bias in favor of major. The conventional view that the expressive quality of the major triad is more "happy" or "satisfied" is also well in accord with considerations of rootedness.

If the root-supporting effect of fb-interval 4 thus has some significance for the practices of tonal music, how about the weaker root supports? Fb-intervals 10 and 2 (minor seventh and major ninth) cannot occur in consonances under the principle that excludes interval class 2. However, the minor seventh occurs alongside the stronger root supports in the most important dissonant chord in tonal music, the dominant seventh chord. The next conventional enlargement of this chord (in major) is the major ninth chord. Hence, these enlargements utilize root supports in the order of root-supporting weights. As is well known, Schenker vehemently renounced the connection between the seventh in the V^7 and the seventh harmonic—to say nothing of the ninth—and sought to explain all dissonances on a purely linear basis.[96] However, linear functions and sonorous appeal do not rule each other out. While dominant sevenths and ninths require linear resolutions to triads in conventional tonality, it seems highly probable that the enormous popularity of such harmonies, and their tendency to govern extended stretches of time in Classic-Romantic music, had something to do with sonorous properties that stem from the connection with the harmonic series.

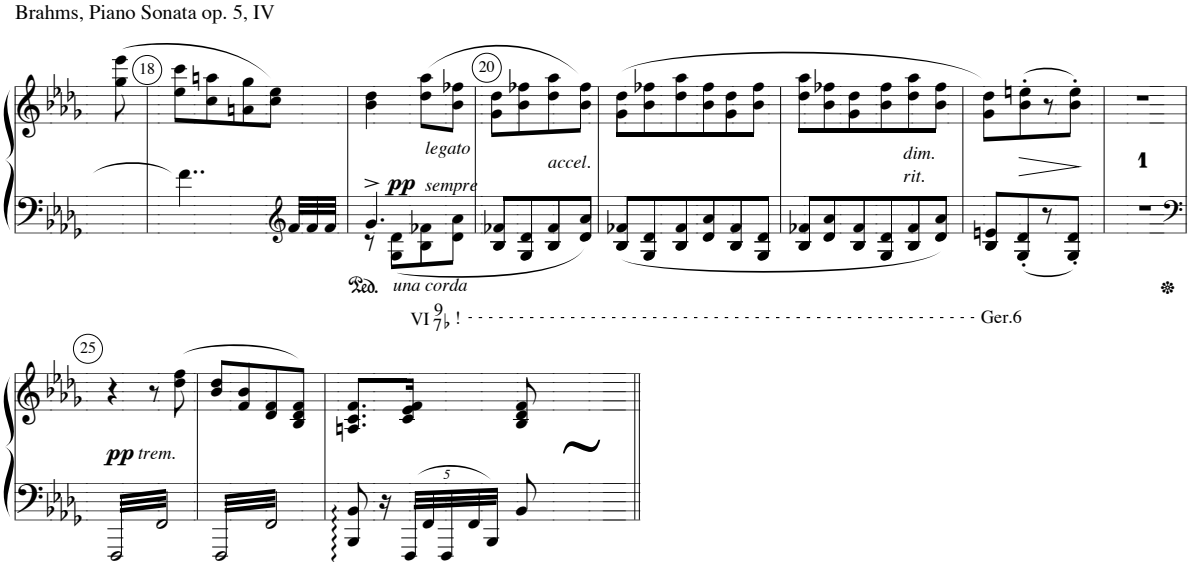
In this connection, one should also mention augmented-sixth chords, whose construction

[96] Schenker 1906/1954 (§ 11) already declares that "no overtone beyond the fifth in the series has any application to our tonal system." Similarly, Schenker 1935/1979 (§ 176) asserts that "where the seventh is not a suspension […] it is a passing tone. As such it has not the slightest relationship to the seventh overtone, which many textbooks hold to be identical with the seventh." These quotes make evident that Schenker rejected the connection between the seventh and the seventh harmonic, but some other passages suggest that he went so far as to reject any kind of significance of the seventh as a vertical relationship. For example, Schenker 1926/1996 (9) states: "Therefore it contradicts the nature of the dissonant passing note to discriminate in any substantial way among the intervals of a fourth, a seventh and a ninth, to say nothing of positing an increasing scale of dissonance for these intervals: *the vertical dimension is altogether excluded*, everything hinges on the horizontal tension alone." (My italics.) While Schenker actually discusses dissonant passing tones in strict counterpoint in this excerpt, the paradigmatic significance of strict counterpoint for his theories permits us to assume that more or less similar considerations apply to free composition. On the other hand, several scholars have observed the inconsistency between Schenker's theoretical statements and analytical practices concerning the position of the V^7; see, for example, Ernst Oster's footnote in Schenker 1935/1979: 64, and Schachter 1999: 201–2.

resembles dominant sevenths (German sixth is enharmonically equivalent with V^7). Like dominant sevenths and ninths, augmented-sixth chords certainly have a linear function but they, too, tend to assume much greater expressive and structural significance than would be characteristic of arbitrary dissonances produced by linear motions.

Whereas Schenker rightfully criticized conventional harmonic analysis for neglecting linear aspects of music, it is also unsatisfactory to go to the other extreme by renouncing the significance of dominant sevenths and ninths and augmented sixths as vertical constructions, just because their syntactic role depends on linear connections. That theorists have been inclined to give special labels to these chords in harmonic analysis reflects the fact that they "superficially" appear to be significant also as verticalities. Important as it is to see through such an appearance and recognize their eventual linear function, "superficial," temporary appearances may also be vital for expressive qualities. Let the particularly eloquent Brahms passage in Example 16 suffice as illustration. Observe how the there is a pedaled pianissimo ninth chord on VI in mm. 19–24 appears as if temporarily ignorant of its linear origins or obligations—or, more hermeneutically, of the agony of the real world. A return to the latter is then brought about in m. 25 by the more or less normal resolution of the German sixth. For understanding the passage, it is essential to recognize both the conventional resolution and the preceding temporary suggestion of sonorous self-sufficiency.[97]

EXAMPLE 16. A Brahms passage exploiting the sonorous qualities of a ninth chord



All in all, the integration of the weaker root supports to harmony, as evident in the examples of **II** and **III**, is not without prehistory in tonal music. Before being integrated into

[97] For an extensive argument for the significance of different temporal perspectives in music analysis, see Lewin 1986.

unconventional referential harmonies, their sonorous features were tried out in non-structural (dissonant) harmonies in conventional tonality.[98] Such an evolutionary process is also evident in the stylistic evolution of individual composers, such as Scriabin. As observed by several scholars (e.g. Baker 1986), the ♭II–V spans with "French sixth" type sonorities, characteristic of his "tonal" period, foreshadow similar structural frameworks in many of his "atonal" works (cf. Example 7c). However, phases in such an evolutionary process are not always strictly chronological. As discussed in **III**, Debussy's mature work, such as the *Préludes*, contains cases in which non-triadic harmonies become referential for entire compositions (such as *Voiles* and *Ce qu'a vu le vent d'ouest*) alongside cases in which similar harmonies assume significance in more limited spans, being subordinate to triads in the overall organization.

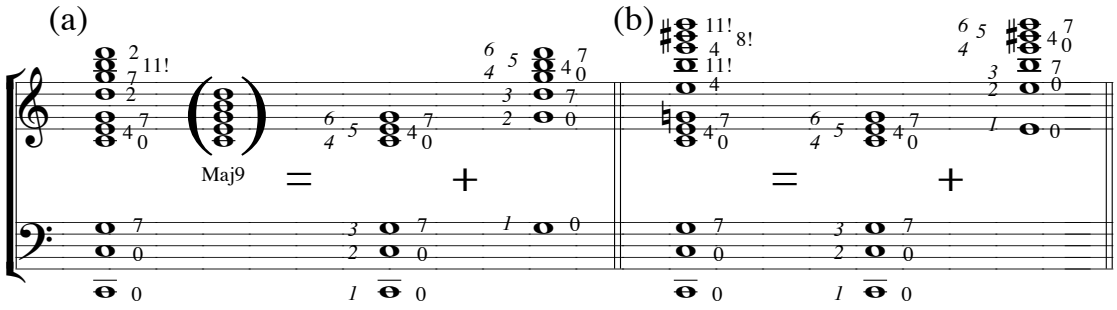
5.2.4. A SPECULATIVE EXCURSION: SECONDARY ROOT SUPPORTS?

The issue of *secondary root supports* is only touched upon in the articles of the present studies (in reference to **II**: Example 7d, **III**: Example 12a,i)—without calling it by that name—but I shall seize the opportunity to elaborate on this issue in connection with the present considerations. By secondary root support, I mean an upper-voice tone that is not a root support in the ordinary sense (i.e., fb-interval 7, 4, 10, 2, or 6), but a root support of a lower upper-voice tone which in turn is a root support of the bass. Consider the two large chords in Example 17a–b. As shown by exclamation marks, fb-interval 11 is present in both chords; in addition, the latter chord (Example 17b) includes fb-interval 8. While 11 and 8 are not root supports in the ordinary sense, they are root supports of root supports, as evident from the sums 11 = 7+4 or 4+7; 8 = 4+4. This is illustrated by conceiving of the two chords as combinations of two major triads: C and G major in Example 17a, C and E major in 17b (as previously, ordinary numbers indicate fb-intervals, whereas italicized numbers indicate corresponding harmonics). As evident from Example 17a, fb-interval 2 is also a root support of root support (2 = 7+7 mod 12); however, since it is also a root support in the ordinary sense, such a secondary relationship holds less interest for the present considerations (at most it could somewhat strengthen the ordinary root-supporting effect).

EXAMPLE 17. The enlargement of the major triad by secondary root supports.

---

[98] One strand in this evolutionary process involves the practice of combining the minor triad with the lower fifth of its conventional root, as demonstrated in Example 14g. The most conventional way to produce this combination, significant in, e.g., Schumann's music, is the superimposition of V and II in major. Later, in music by, e.g., Debussy and Sibelius, the combination occurs frequently as a superimposition of IV (with major third) and I in minor (see, for example, the passage from Debussy's *La sérénade interrompue* in **III**: Example 12b,ii)—or in less conventional harmonic practices that allude to such a combination (see, for example, the analysis of a passage in Sibelius's Symphony no. 7 in Väisälä 2002, in reference to Example 7).



The notion of secondary root support relies on the following kind of reasoning. Let pitch r be a root support of q and q a root support of p. This means that r has a tendency to belong to a pattern governed by q, which in turn has a tendency to belong to a pattern governed by p. Realizing these tendencies produces a two-level hierarchy, in which p governs q and q governs r. Hence p governs the overall hierarchy, including r; this is the meaning of calling r a secondary root support of p.[99] This notion involves a cognitive extension, which extends a psychoacoustically influenced organizational feature of certain intervals to apply to combinations of such intervals.

For evaluating the cogency of such a notion, some kind of background is given by observing that a comparable extension is evident in the principle of octave equivalence (or relatedness). Pitches separated by multiple octaves are functionally octave equivalent even though the psychoacoustical properties supporting this relationship only apply to simple octaves (cf. section 2.2.1). Another point of comparison is given by the circle of fifths. The conventional notion of the circle of fifths as some kind of a measure of harmonic relatedness in tonal music may have psychoacoustical justification in the properties of the fifth, but conceiving of other intervals as combinations of fifths is a cognitive extension. However, there is one respect in which these concepts differ from secondary root supports. The relationship between tones separated by multiple octaves does not require that the intervening octaves are actually present, which reflects the strength in which the principle is cognitively established. Similarly, the circle of fifths is sometimes used as a measure of harmonic relatedness between harmonies without presupposing that the intervening elements are actually present in music.[100] By contrast, I shall confine the notion of secondary root supports to cases in which the intermediate element(s) in the sum construction is (are) actually present. For example, for fb-interval 11 to function as a secondary root support, either 7 or 4, or both, must be concretely present in the harmony to support the perception of 11 in terms of the sum 7+4 or 4+7.[101]

---

[99] The significance of such secondary relationships is also observed by Vincent Persichetti (1962: 24): "Resonant harmony is not formed by seeking higher and higher overtones but by using overtones of overtones."

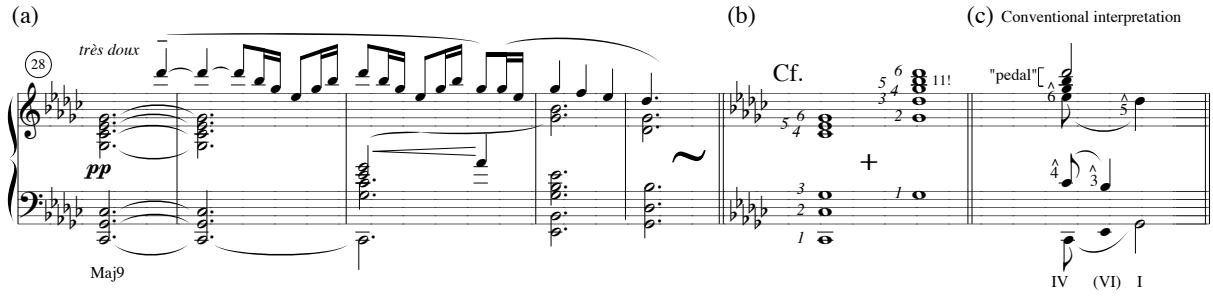
[100] Whether and to which extent the circle of fifths actually pertains to musical organization and cognition is a complex issue I will not consider here.

[101] Very speculatively, however, we could consider what would follow from lifting this constraint. As the present discussion makes evident, the strongest candidate for a secondary root support is fb-interval 11, the sum

In the present studies, the idea of secondary root supports is only invoked in connection with fb-interval 11. The Debussy analyses in **III** highlight the chord in Example 17a, called the *major ninth chord* (a more compact spacing is shown in parentheses). Another significant occurrence of fb-interval 11 is at the end of Scriabin's *Vers la flamme* (**II**: Example 7d).

In the Debussy analyses, this notion helps to explain the particular expressive and structural significance of bass of the major ninth chord in relation to the upper voices. Consider the "minor seventh chord" formed by the upper voices (E–G–B–D in Example 17a). As illustrated in Example 14k above, there are four candidates for roots, each of which is supported by three tones of the "minor seventh chord" (no bass tone is supported by all four tones). The "minor seventh chord" occurs in several pieces in the first book of *Préludes*, most often as the black-key transposition D♯–F♯–A♯–C♯ (or E♭–G♭–B♭–D♭). For this transposition, the four root candidates are F♯, D♯, B, and G♯. The first three are employed in several *Préludes* (**III**).[102] In *Ce qu'a vu le vent d'ouest*, B, D♯, and F♯ form the structural bass arpeggiation. In *La fille aux cheveux de lin*, a similar arpeggiation (notated with flats) occurs at the foreground; see Example 18a (Example 18 reproduces Examples 10c and 12b,i from **III**). While the significance of these three bass notes can be explained on the basis of ordinary root supports, the secondary root-supporting status fb-interval 11, which permits all the upper voices to participate in the root-supporting pattern, helps to account for the special kind of "satisfactory"

EXAMPLE 18. The major ninth chord in Debussy's *La fille aux cheveux de lin*, mm. 28–32.



effect produced by the B/C♭, an effect very much evident in the *La fille* passage. (From the structural perspective, the major ninth chord on C♭ does not function as the referential harmony—Example 18c presents a structural interpretation based on conventional tonality—but it is a kind of expressive culmination in the piece.)[103]

---

of the two strongest root supports (7+4). On the other hand, its inversion, 1, is the sum (mod 12) of the two "root-opposing" fb-intervals (5+8). If we were to assume that these sum constructions affect the cognitive properties of 11 and 1 even when the intermediate element is not present, this would produce additional support for the consonance–dissonance relationship between these intervals that is characteristic of some music by Schoenberg and that is elsewhere here (sections 3.2.3.4. and 5.1; **I**: 237) explained on the basis of the partial octave generalization of the proximity principle.

[102] A possible reason why G♯ is less important is that it forms the "root-opposing" fb-interval 5 with C♯.

[103] The central significance of the major ninth chord in *La fille* is also supported by an early sketch showing the

In **I–III**, no cases are discussed in which fb-interval 8 could be conceived as a secondary root support (as 4+4), as illustrated in Example 17b. However, Debussy's *Préludes* contain a passage for which such a conception seems very pertinent: the conclusion of *Ondine* (second book of *Préludes*, no. 8); see Example 19. The texture consists of alternating D major and F♯ major triads. It is beyond the scope of the present discussion to offer a precise structural interpretation of the details in this alternation, for example, whether and to which extent the D–C♯ and A♮–A♯ relationships between the two triads function as neighboring motion. Suffice it to observe a couple of features suggesting that that there is at least a hint of C♯ and A♯ (fb-intervals 11 and 8) participating in some kind of an overall harmony on D. First, the F♯-major block chords in mm. 66 and 68 occur literally above the D. Second, the A♯6 and C♯7 in these chords, repeated for the last time in measure 72, are never followed by A♮ or D in the same register; hence, by virtue of the proximity principle of voice leading, they are in some sense left "hanging." While there are thus hints to a kind of coexistence of the D major and F♯ major triads in harmony, these triads are also consistently separated by textural and registral means. This reinforces the conception of A♯ and C♯ as primarily related with F♯; they are heard as root supports of F♯, which in turn supports the bass D. The textural and registral features seem ideally arranged to support the conception of C♯ and A♯ as secondary root supports.

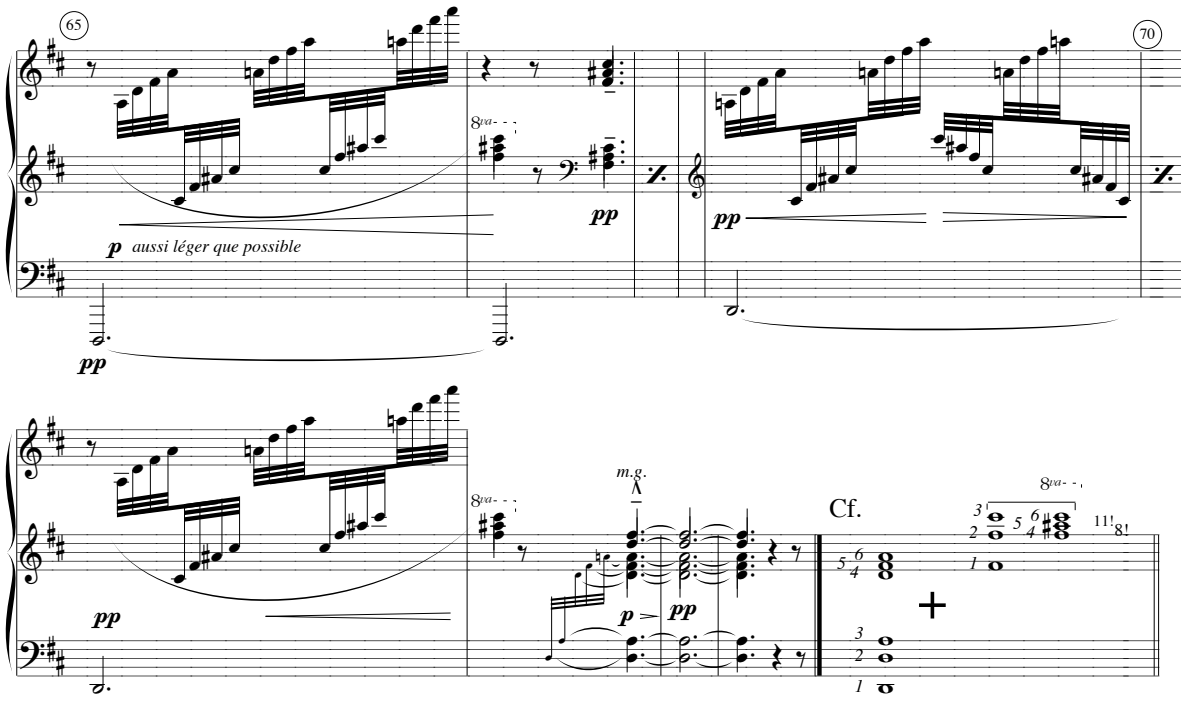
The conclusion of *Ondine* relates with preceding motivic features in a way whose beauty warrants a brief digression. Example 20 offers illustration. Notational symbols are introduced in Example 20a. Motive x (A–B♭–C♯) and its inversion, and the transpositions of the two at the lower fifth, may be understood primarily in terms of pitch-class sets (since only one transposition is important, this is denoted simply as Tx). Both these three-note motives and their combinations, including the overall eight-note pitch-class set, are motivically significant. Example 20b shows how the overall bass line relates with these motives (without specifying the structural relationships within the bass line).[104] Example 20c illustrates important foreground features. A remarkable feature in the conclusion is the way in which registral and textural manipulation of familiar motivic material gives rise to new sonorous qualities. Initially, the semitone (or, more precisely, ic-1) relationships in motives x and Ix (A–B♭ and C♯–D) have a rather pungent character, which is emphasized by gestural abruptness (see especially m. 4).

---

piece as beginning with that chord (**III**: Example 10d).

[104] The connection between the overall bass line and the rapid figure in mm. 11 is also observed by Claudia Zenck-Maurer (1974: 98). Also Friedmann 1982 contains some discussion of the motivic relationships in *Ondine*.

Example 19. Debussy, *Ondine*. Conclusion.

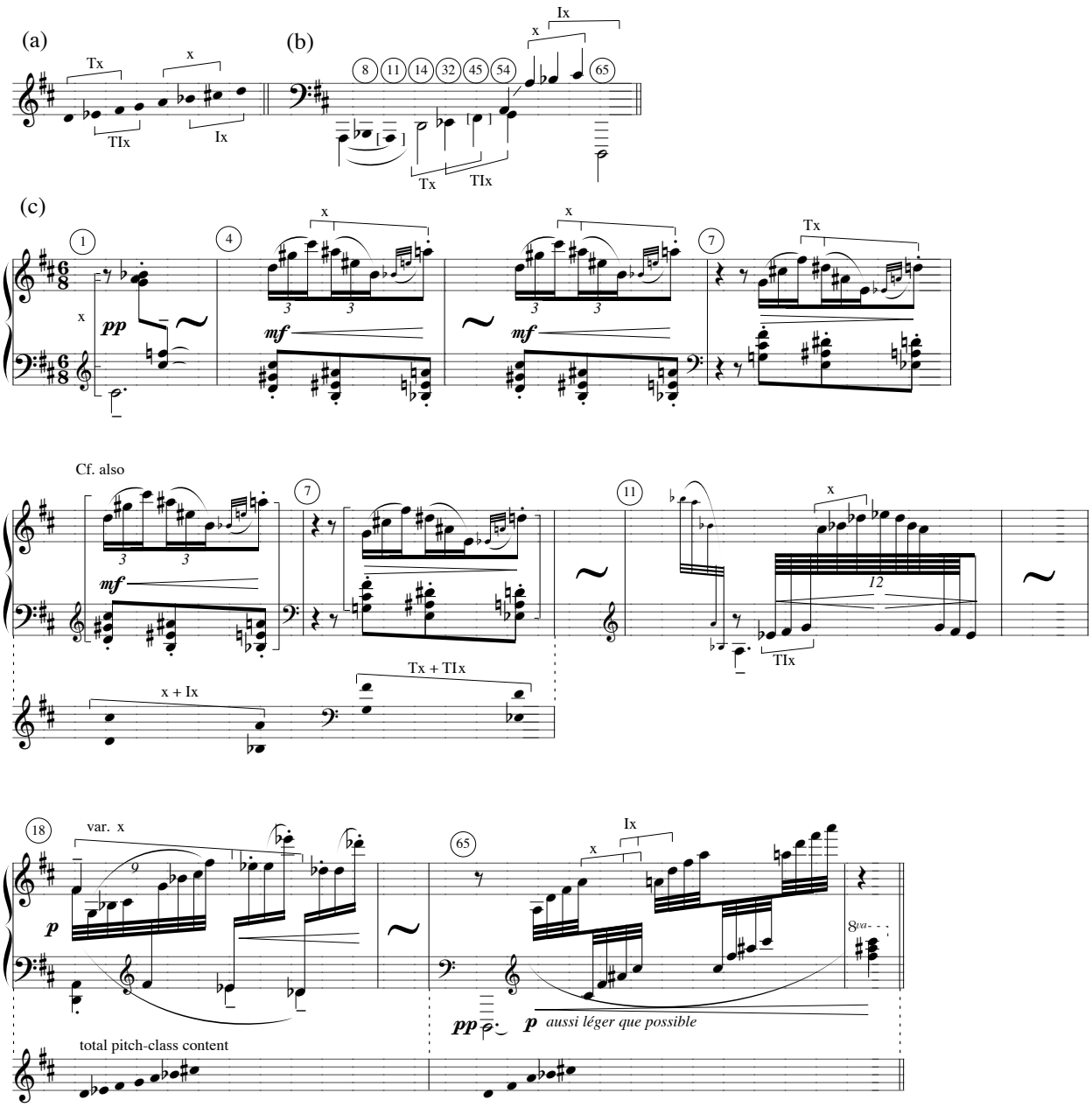


Subsuming A♯/B♭ and C♯ into the F♯ major triad in the conclusion alleviates this pungency by breaking up the semitones, and gives these tones a more compliant character as secondary root supports in the overall pattern.

11 (7+4 or 4+7) and 8 (4+4) do not, of course, exhaust the possibilities to express fb-intervals as sums of root supports. However, given that the entire discussion of secondary root supports may be regarded as speculative, it seems sensible to confine these speculations to sums that most probably pertain to our perception and cognition. Factors relevant to this issue include the root-supporting weights of the addends and the number of them. Regarding the first factor, it may be sufficient to confine our considerations to the two strongest root supports, 7 and 4. For sums consisting of two addends, there are no other possibilities than those already discussed. By increasing the number of addends, we obtain 7+7+7 = 9 (mod 12) as the next most probable extension to candidates for secondary root supports. In more conventional terms, this means that in a chord of three superimposed fifths, the major sixth between the highest and lowest might function as a secondary root support. For an example of such a harmony, one may consider the conclusion of Debussy's *Reflets dans l'eau*, as illustrated in **III**: Example 15a.[105]

[105] One factor relevant for the issue of secondary root supports is whether the sum interval is related to rootedness in a more direct way, so as to overshadow the significance of a secondary relationship. For example, consider the sum 7+10 = 5 (mod 12). Even if we were to approve 10 (the third strongest root support) as an addend, the significance of such a sum would be diminished by the strong "root-opposing" quality of fb-interval 5 (fourth). As regards the present examples of secondary root supports, 11 (7+4 or 4+7) and 9 (7+7+7) hold no such direct relationship with rootedness, whereas 8 (4+4) is "root-opposing." Consequently, the latter sum construction requires stronger contextual support in order to override the direct relationship.

EXAMPLE 20. Debussy, *Ondine*. Motivic organization



Apropos contextual support, in addition to particular contexts that enhance the pertinence of a sum construction to a musical passage (as in Example 19), certain sum constructions may also be supported by practices in wider cultural contexts, and by the influence of these practices on our cognition of intervals. For example, conceiving of 9 as 7+7+7 (mod 12) might be supported by the conventional notion of the chain of fifths and its relevance for cognition. The conception of 11 as 7+4 is also strongly manifest in the conventional tonal system, in the relationship between the tonic and the leading tone. The leading tone is typically harmonized as the third of the V scale degree. Hence the conception of ro-interval 11 (major seventh) as a sum of 7 (fifth) and 4 (major third) is also supported by our cognitive familiarity of such a sum construction in tonal music, even though the occurrence of the elements involved in this sum within one functionally consonant harmony is a non-conventional feature.

5.3 STREAMING AND THE PROXIMITY PRINCIPLE OF VOICE LEADING

The connection between streaming and the proximity principle of voice leading—the main principle for satisfying the harmony/voice-leading condition—was introduced in section 2.2.1. Streaming refers to the process in which our hearing groups information into simultaneous streams. Bregman 1990 may be considered a classic general presentation of the research on streaming, its underlying factors, and some of its musical implications. Of the underlying factors, pitch distance is relevant to the present discussion. A large interval between two pitches enhances their tendency to be segregated into two streams; the correspondence of this phenomenon with the proximity principle of voice leading is obvious.

It should be added, however, that other factors of streaming, such as timbre, may have musical implications that interfere with the pitch-based proximity principle. In the two songs analyzed in the present studies (Berg and Webern; see **II**), the timbral difference between the piano and voice parts tends to diminish the structural significance of the proximity principle of voice leading.[106] Since most works in the present repertoire are written for piano, this phenomenon is not evident in them. Generally speaking, however, one should be aware of the potential significance of timbre for structural relationships when considering post-tonal prolongation.

Despite the relationship between streaming and voice leading, the two cannot be equated. The latter involves extensive cognitive generalizations and extensions, concerning the dimensions of both pitch and time. Temporal extensions have already been discussed in section 2.2. As for pitch, the "default" borderline between voice-leading intervals and arpeggiations, situated between two and three semitones (section 3.2.3.4), by no means always corresponds with the pitch distance inducing stream segregation. This pitch distance depends on the listener's conscious efforts. The lower limit of hearing two streams is called the *fission boundary* whereas the upper limit of hearing one stream is called *temporal coherence boundary*.[107] The latter boundary depends on the temporal separation of tones to a much greater extent than the former. What is significant for the present considerations is that the fission boundary approximately agrees with the "default" borderline in the proximity principle of voice leading. That is to say, listeners are able, at will, to hear two alternating pitches as forming separate streams unless their distance goes under three our four semitones (depending somewhat on the temporal rate). The "default" borderline may thus be understood as based on a cognitive generalization of the fission boundary.

The "default" borderline also agrees with the critical bandwidth, which, in turn, links with

---

[106] In fact, timbral difference also tends to diminish the significance of the proximity principle of spacing (**II**: note 32).

[107] These concepts were introduced by L. P. A. S. van Noorden in his unpublished dissertation "Temporal Coherence in the Perception of Tone Sequences" (Eindhoven University of Technology, 1975). My present discussion is based on Bregman's (1990: 58 ff.) report of van Noorden's research.

the proximity principle of spacing (section 5.1). As discussed in section 3.2.3.4, the correspondence between the two aspects of the proximity principle clarifies musical organization. Such clarification may partially explain the success of musical systems in which these aspects correspond. In addition, some authors have also hypothesized that the critical band may bear a direct link with streaming.[108] If this is the case, the two aspects of the proximity principle have a common background factor in the critical band.

[108] For example, Bregman (ibid.: 407) observes that "[i]t may not be a coincidence that the value of the three or four semitones [of the fission boundary] is about the same as the width of the critical band […]."

6. CONCLUDING REMARKS ON THEORETICAL PRINCIPLES AND ANALYTICAL RESULTS

Above, I have discussed the concept of prolongation on the basis of Straus's four conditions (section 3). I have described the ways in which the conditions concern two basic requirements of organization: consistent norms of harmonic stability and melodic relationships. While the requirement for such functional norms is shared by tonal and post-tonal prolongation, the norms are different. Two sources for such norms have been identified: the referential harmony and the proximity principle of voice leading (see especially section 3.2.3.4). The differences between tonal and post-tonal functional norms stem primarily from the former source.

I have also discussed how one should conceive of harmonies and intervals in prolongational studies (section 4), and how psychoacoustical factors may influence functional norms (section 5). Such influence is manifest both in the formation of the referential harmony and directly in the proximity principle.

In this concluding section, I take a closer look at the concept of referential harmony, as evident in the present studies (section 6.1), and summarize the principles for meeting the four conditions (section 6.2). This discussion is followed by some considerations on the analytical results (section 6.3).

6.1 THE REFERENTIAL HARMONY

6.1.1 TRACING THE REFERENTIAL HARMONY; SOME HINTS

In conventional tonal prolongation, the status of the referential harmony, the tonic triad, involves two main aspects. The tonic triad is the normative stable surface harmony and offers the basis for the linear overall structure. The present analyses are based on the search for a referential harmony that would, as far as possible, fulfill these two functions.

I will not attempt to describe a generally applicable strategy for tracing the referential harmony, but only make some observations of where the clearest clues to the referential harmony can be found in the present examples.[109]

[109] These observations do not purport to describe my actual processes of tracing the referential harmony. Such processes are more complex than can be described in this way; moreover, they have largely fallen from memory.

Since prolonged entities normally appear at either end of the prolongational span, beginnings and conclusions of pieces are logical points for starting the search for clues to the referential harmony. However, the choice between these alternatives may not be easy; moreover, it is possible that the beginning and/or the end fall outside the main prolongational span, functioning as a non-structural introductory or coda-like attachment. Determining which tones actually belong to a harmony may also cause two kinds of difficulties. First, it may be difficult to distinguish harmonic from non-harmonic tones within a given temporal span. Second, it may be difficult to determine the temporal span pertinent to a presentation of a harmony. As discussed in connection with *Vers la flamme* and *Voiles* in section 3.1.2, the presentation of a complete harmony may take considerable length.

In a couple of cases, a clue to the referential harmony is given by a clearly expressed harmony at the *conclusion* of the composition. In the analyses of Schoenberg's op. 19/2 (**I**) and Debussy's *Ce qu'a vu le vent d'ouest* (**III**), I start my discussion from such a clue. In the Debussy, the concluding harmony is an emphatically repeated block chord. In the Schoenberg, it is not a block chord (it consists of the entire pitch content of the final measure), but perceiving it requires only a minimal amount of temporal integration and is facilitated by clear registral means.[110] In both cases, the concluding harmony creates a kind of "stabilizing" effect, and the assumption of its referential status is rewarded by important analytical findings. However, these findings also give feedback that ultimately requires the precise identity of the referential harmony to be specified. Hence, in the Schoenberg analysis, I ultimately identify the *opening* harmony as the primary transposition of the referential harmony, owing to the "inclusion factor" discussed in section 3.2.3.5: the lower-voice events compose out this chord rather than the concluding one. In the Debussy, the conception of the referential harmony is specified by the observation that the lower fifth of the bass of the concluding harmony enlarges the harmony in the overall structure to form a "major ninth chord"; this issue will be discussed in subsection 6.1.2 in reference to Example 23.

Simple block-chord *openings* are found in Berg's op. 2/2 and in Scriabin's *Vers la flamme*. As in the cases discussed above, these chords offer clues to the referential harmony but are not quite identical with our ultimate conception of it. As demonstrated in Examples 7c and 10b above, both cases open by chord denoted as P, which is later enlarged to U by arpeggiation. In both pieces there is additional complexity in temporal relationships. In the Berg, the structure is completed by a salient block-chord U (m. 15), but this is not the end of the piece. It is followed by a passage that I interpret as a transition to the next song in the opus (**II**: Example

---

They involve both instinctive familiarization with the music and more or less conscious efforts of trial and error.

[110] Additional motivation for starting the discussion from the concluding harmony is given by the desire to relate the considerations in **I** with those by Travis (1966), who sees the concluding harmony as the "tonic sonority"—a notion corresponding to the present "referential harmony"—, and by Straus (1987), who criticizes Travis's analysis on the basis of the set-theoretical properties of this harmony.

11). In the Scriabin, the overall structure is based on several locally referential harmonies, the first of which is U. In the mutual hierarchy between these harmonies, a chord called A^{+9}—which consists of all the root supports and an "added" fb-interval 9—is identified as the primary or overall referential harmony. This choice is justified on the following grounds: First, it arrives when the structural arpeggiation E2–B1–E1 reaches its goal (m. 95). Second, the other harmonies are easily relatable to it (by semitone shifts, inclusion relationships, and chord inversion).[111] Third, its rootedness is superior among the harmonies.[112] While the last observation is by no means decisive—rootedness does not always correlate with structural weight—it supports the other arguments and calls attention to the relevance of rootedness for Scriabin's organization. In the temporal organization, however, A^{+9} is not the final harmony. At the very end, fb-interval 10 is replaced by 11, whereas other fb-intervals remain stationary (on E1). The structural function of the concluding harmony may be compared with the Picardy third: the substitution of 11 for 10 highlights the conclusion but is a purely local phenomenon with no ramifications for the structure.

In Debussy's *Voiles* and Webern's op. 3/1, the clearest clues to the referential harmony are given not by block chords, but by melodic motions. I begin my analyses (**II**) by discussing how harmonic frameworks may be traced by interpreting such motions at the opening of the pieces on the basis of the proximity principle of voice leading, allowing for rhythmic and motivic relationships or "embellishment clues" (section 3.2.3.3). These interpretations are later corroborated by some vertical occurrences of the horizontalized harmonies, but such occurrences are less salient than in the Scriabin and Berg examples.

Some special difficulties in the Debussy and Webern examples may be worth mentioning. In the opening of *Voiles*, the relationships between the melodic material and the governing chord (Q; see Example 4b) are relatively straightforward (for a foreground graph, see **II**: 16),[113] but this governing harmony does not contain all tones in the referential harmony. It is enlarged by one tone (D) through the cadential gesture at the end of the opening section (to form chord U; see Example 4b–d); this issue will be taken up in the next subsection.

In Webern's op. 3/1 the most notable difficulty concerns the separation of harmonic and non-harmonic tones. The clues given by the vocal part lead us to a notion of referential harmony

---

[111] In terms of fb-intervals, the opening U is a subset of A^{+9}. However, the relationship between these harmonies also involves a semitone shift since fb-interval 6 ($A\sharp/B\flat$) in the second lowest voice of T_0U is replaced by 7 (B) in T_0A^{+9}. Intervening between U and A^{+9} in the overall organization is a "7-inversion" of the latter (an inversion in which fb-interval 7 of the root-position chord lies at the bass). For illustration, see the present Example 22a(v) and **II**: Examples 7–9.

[112] A^{+9} is more rooted than U owing to the substitution of the strongest root support 7 for 6 in the second lowest voice. Incidentally, this substitution also diminishes roughness (cf. section 5.1).

[113] Adele Katz (1945: Example 93) interpreted the opening melodic material of *Voiles* as based on the horizontalization of the "augmented triad" that forms the upper voices of chord Q. Hence, even without a common theoretical background of post-tonal prolongation, my analysis and Katz's are essentially similar, reflecting the straightforward way in which the melodic motions relate to the governing harmony (Katz's analysis came to my attention only after I had made the basic observations in my analysis).

which permits a coherent description of the structure. However, under this description the piano part includes several non-harmonic tones whose functional status is not clearly evident in their outward character. In other words, these tones may be interpreted as forming non-harmonic embellishments, but such an interpretation is not supported by strong "embellishment clues." Moreover, while I think this interpretation is "correct" with respect to prolongational considerations, it does not capture all significant aspects of harmonic organization. Though the structurally primary aspect of these tones is their non-harmonic, linear relationship with the referential harmony, one may also identify principles regulating their vertical relationships with other tones; see the discussion of "supplementary principles of harmony" in the analysis in **II**. This notion—that harmonic organization does not emerge solely from the referential harmony but may involve significant supplementary aspects—would seem to have considerable potential significance for studies of post-tonal prolongation; it permits a greater variation of foreground harmonies while not rejecting the notion of the referential harmony. It should be noted that "supplementary principles of harmony" are also not foreign to conventional tonality; this kind of issue was considered in section 5.2.3. If the concept of referential harmony is inadequate for capturing all significant aspects of harmony in the Webern example, the same can be said of the triad with respect to passages such as mm. 19–24 in Example 16.

6.1.2 SUBSETS AND THEIR ENLARGEMENT IN THE PRESENTATION OF REFERENTIAL HARMONIES

As evident from the preceding discussion, several of the present examples start with a subset of the referential harmony, which is subsequently enlarged or complemented by arpeggiation to form the actual referential harmony. Such enlargement was already discussed in section 3.1 as was the literal sense in which the subsets have to be understood. They cannot be regarded as implicitly completed, in the manner of incomplete triads in conventional tonality, prior to their literal completion. However, some important questions relevant to such enlargement remain to be considered. First, what is actually required to justify reading a horizontal interval as arpeggiation—in other words, as standing "basically" for a verticality? Furthermore, how does one distinguish between consonant and dissonant arpeggiations? Second, if the referential or governing harmony is in its entirety formed only after considerable temporal delay, how does this affect its relevance to the listener's experience? In the following, I illuminate these questions on the basis of some examples.

Regarding the first question, the crucial role of the proximity principle of voice leading for the interpretation of arpeggiation is evident from all the preceding discussion. However, this principle is not alone sufficient: by no means do all "leaps" arpeggiate an interval within one harmony. The application of the proximity principle, combined with "embellishment clues,"

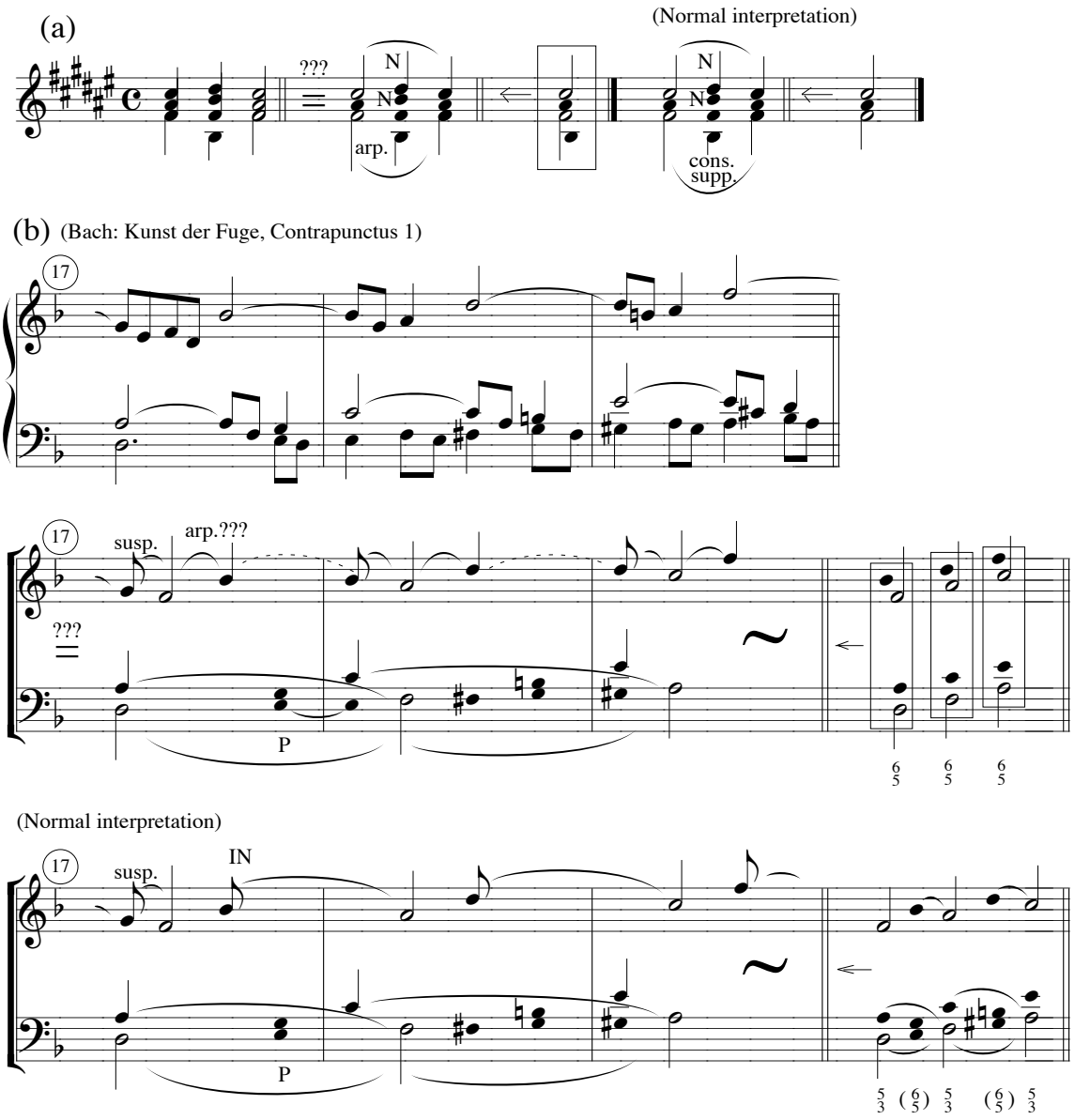
may lead to the revelation of a harmonic framework (see the preceding subsection), but it may also lead astray. This can be demonstrated by considering the absurd results such an analytical strategy could produce in progressions of conventional tonal music. Examples 21a and 21b show two "readings" according to which triads are enlarged by arpeggiation, forming "governing harmonies" shown in boxes. While these examples are absurd indeed, they are logically supported by the proximity principle of voice leading and they do not contradict any "embellishment clues." Someone with no previous understanding of conventional tonal syntax—a "Martian theorist," so to say—could not be reproached for making such readings as first guesses of how the music works. The only fault in these readings is that interpreting the leaps as arpeggiation lacks here all substantiation in the harmonic language of conventional tonality. The "chords" inside the boxes do not occur as relatively stable entities in tonal music. While the readings are sufficiently supported by the norms of melodic relationships (conditions #3 and #4), they totally ignore norms of harmonic stability (conditions #1 and #2). These readings also lack explanatory power. They cannot tell us anything of how to relate these passages in a meaningful way to other phenomena in a tonal context or why it is just these harmonies (those in the boxes) that have been composed out. The availability of better explanations is a definite reason to abandon them.

While these points may be more or less self-evident—for "Earthling theorists," that is—explicating principles that underlie prolongational analysis in tonal music helps us, once again, to clarify corresponding principles in post-tonal music. Example 22a reproduces the reading of the opening of *Vers la flamme* from Example 7c, adding some detail in graph v. The sequential progression and the ascending leaps in the top voice bear some outward resemblance with the Bach passage in Example 21b. What justification is there to claim that this time the leaps, from fb-interval 10 to 2 (D–F♯ at the opening), actually conduct arpeggiation and enlarge the opening harmony from U to P?

As already observed, the U chord appears as a verticality in m. 19 (at T_6), and fb-interval 2 in its top-voice (C) fulfills, after a registral transfer, an important linear function in m. 23. Hence the function of the opening D4–F4♯ interval as arpeggiation within chord U is concretized and substantiated by the subsequent events. Nevertheless, these arguments do not yet show that the U chord governs through the first four measures, as indicated in Example 22a. It could be subordinate to P, as a "privileged" dissonant chord, as indicated in Example 22b. By "privileged" dissonance I mean a chord comparable to a dominant seventh in tonal music, a chord with a clear harmonic identity but whose syntactic function ultimately depends on the resolution of the dissonance.[114] In Example 22b, the occurrences of fb-interval 2 in

[114] Although the concept of "privileged" dissonances is foreign to Schenker's theory (see note 96 above), the privileged status of V^7 is clearly manifest in the organization of conventional tonality and, one might add, in Schenker's analytical practice. For a radical example relevant to the present considerations of temporal integration, consider the reading of Haydn's "Chorale St. Antonii" (Schenker 1935/1979: Fig. 42,2). The V^7

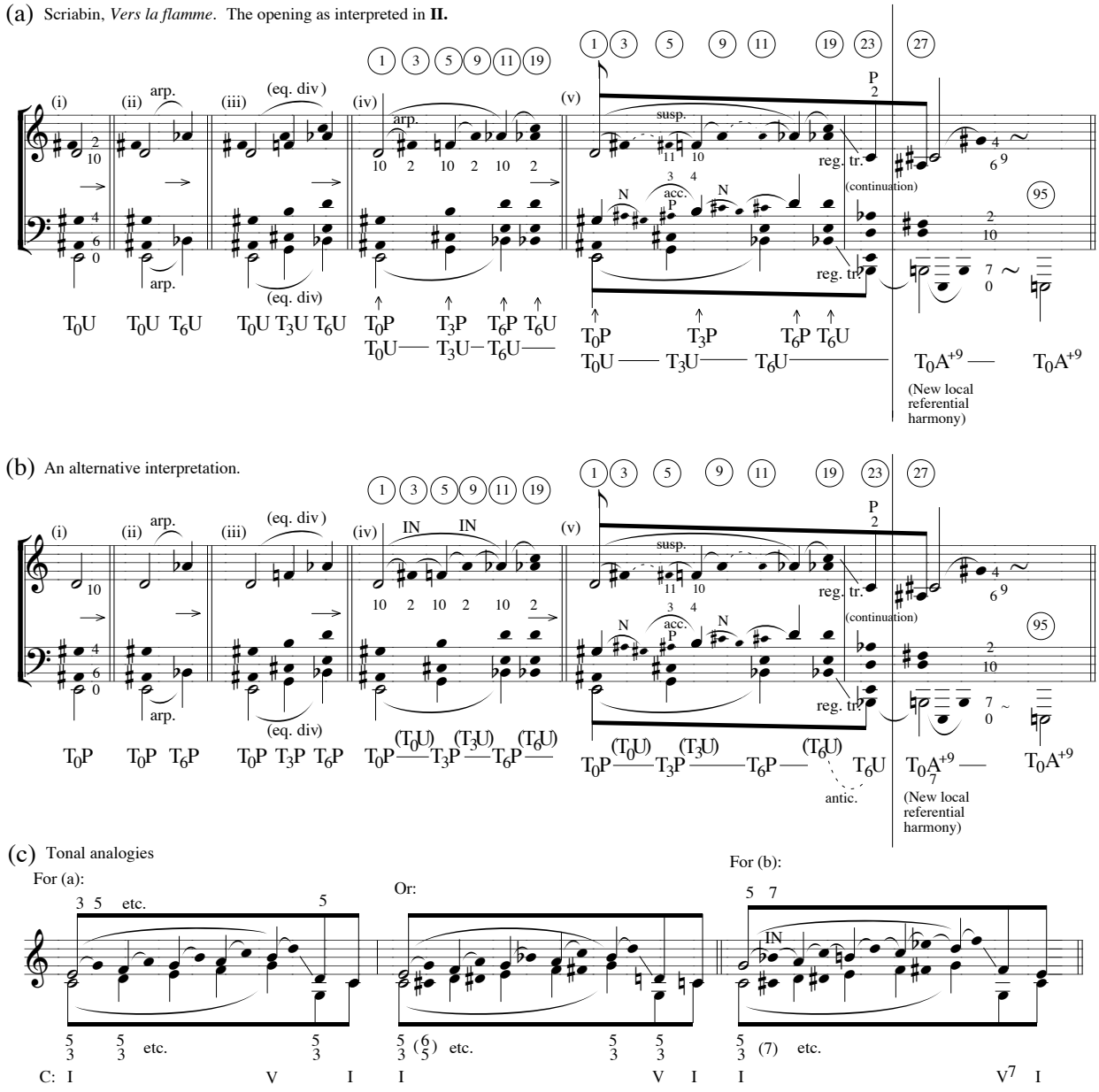
EXAMPLE 21. Absurd readings of chord enlargement in tonal music



mm. 3 (F♯) and 9 (A) are understood as incomplete neighbors that resolve by a descending semitone. Even the use of fb-interval 2 as a middleground passing tone (m. 23) does not really testify against its dissonance; a comparable function can be fulfilled by a dominant seventh in conventional tonality. As shown in Example 22c, tonal analogies may be construed for both of

governing mm. 11–18 is formed through the temporal integration of the V in mm. 11–13 and 18, and the top-voice $\hat{4}$ of the IV^6 in m. 15, without never materializing at the surface. It seems unlikely that any other dissonant sonority could have a governing position in similar circumstances.

EXAMPLE 22. Alternative readings of the opening of Scriabin's *Vers la flamme*. Tonal analogies.



the alternative readings, suggesting that neither of them is implausible. The crucial question is whether the top-voice arpeggiation from fb-interval 10 to 2 connects two consonances or proceeds from a consonance to a dissonance.

It may be noted that the pictures that Examples 22a and 22b give of the organization are actually not too far from each other. According to the former reading, fb-interval 2 is consonant; however, it is less strongly established than the other consonances (0, 6, 10, 4). According to the latter reading, fb-interval 2 is dissonant; however, it has "privileged" status, in comparison to dissonances such as the 3 and 11 at the beginning of measure 5. The order of interval stability is the same in both readings. In fact, the material within the opening section (mm. 1–26) does

not seem to provide enough evidence to choosing between these alternatives. The circumstances would be similar if we were to study the tonal progressions in Example 22c with no previous knowledge of the tonal consonance–dissonance system, i.e., merely on the basis of what takes place in these progressions. There is nothing from which to infer that the fifth is consonant but the seventh dissonant. For gaining this knowledge, we would have to widen our perspective beyond the progression itself and study the larger context of "common practice" tonality.

Widening interpretational perspective is also required for the Scriabin. As already observed, the harmony governing the opening section—be it T_0U or T_0P—is only locally referential. It is followed by a 7-inverted T_0A^{+9} in m. 27 and finally by the primary harmony in the overall organization, the root-position T_0A^{+9}. (See the notations in Example 22a–b. 7-inversion is indicated by "7" beneath "A"; in a 7-invertion, fb-interval 7 of the root-position chord functions as the bass. Interval numbers are indicated in mm. 27 ff. in relation to the root E instead of the bass B; they are not literal fb-intervals.)[115] During the spans governed by these harmonies, fb-interval 2—or, when the 7-inversion prevails, the corresponding interval in relation to the root—does not show any clear dependency on resolution. Considerations on the general tendencies in Scriabin's later output also support the view of fb-interval 2 (the major ninth) as an integral, "consonant" element of harmony. Hence relating the opening of *Vers la flamme* to larger contexts offers support to Example 22a rather than 22b.

Psychoacoustical considerations provide additional illumination. Fb-interval 2 ("major ninth") supplements the pattern of root supports formed by other tones in the harmony.[116] Fb-intervals 10 and 2, which occur between the outer voices in the *Vers la flamme* passage, are about equally strong root supports. The perceptual similarity of these intervals is enhanced by their interval-class identity; moreover, there is no essential difference in the critical-band effects produced by them. By comparison, the fifth and the seventh, occurring in the rightmost passage of Example 22c, bear a much stronger perceptual contrast. The fifth (fb-interval 7) is a much stronger root support than the minor seventh (10). Even more importantly, the seventh violates the crucially important principle of avoiding octave generalizations of "stepwise" intervals in consonant harmonies of conventional tonality ("the octave-generalized proximity principle of spacing"; see section 5.1); the seventh also produces much greater critical-band effects than the fifth. Hence the functional contrast between the fifth and the seventh in conventional tonality is underlined and supported by strong perceptual differences, whereas no similar support is found for the reading in Example 22b in which fb-intervals 10 and 2 hold contrasting functional status.

---

[115] The root E is actually present in m. 28, but is structurally subordinate at this point to B; see the discussion in reference to Example 8 in **II**.

[116] The low registral position of fb-interval 6 deviates from this pattern, but this is "corrected" by fb-interval 7, which substitutes for 6 in the primary A^{+9} chord. This change is highlighted by the reharmonization of the opening subject in m. 107 ff.; see **II**: Example 7.

While these considerations exemplify the kind of justification required by the readings based on the enlargement of harmonies, they do not illuminate the second of the above questions, concerning the effects of such enlargements on the listener's temporal experience. Even though chord U, which includes fb-interval 2, structurally governs the first four measures in *Vers la flamme*, the presence of fb-interval 2 is not felt until m. 3. One may say that for "competent listeners" of Scriabin (or of music showing related practices),[117] this enlargement is not a great surprise—it has the character of pattern completion—but it would be exaggerated to claim that fb-interval 2 is implicitly present in mm. 1–2 in a way similar to the implied tones in conventional tonality. Hence the temporal span is governed by a structural entity whose identity becomes clear only after a temporal delay.

In conventional tonality, similar delays are not manifest in the construction of individual harmonies, but somewhat analogous processes are manifest in other aspects of organization. As discussed in section 3.1.2, the structurally governing entity may appear at the end of a progression (in the ES schema). For the present considerations, an especially interesting point of comparison is given by pieces that open with or are based on an auxiliary cadence. In such pieces, the primary transpositional level of the referential harmony is revealed after a delay, whereas in the present cases the delay concerns the construction of the harmony. In both cases, the experiential effect depends naturally on the length of the delay. If the delay is short—like the two measures in the Scriabin—the process is easily perceptible with the help of short-term memory, making the relationship between the interpretation and the musical experience relatively unproblematic.

Longer delays—like the twenty measures in *Voiles* (Example 4b–c)—bring about more noteworthy problems. If there is a long period during which the construction or the primary transpositional level of the referential harmony is not manifest, there would seem to be a less concrete sense in which it is "referential." Consider the not infrequent practice in Romantic music to begin and end compositions in different keys. Such compositions are sometimes interpreted in terms of large auxiliary cadences. However, the degree of correspondence of such interpretations with the musical experience depends on the extent to which there are features buttressing the primary status of the concluding tonality, features indicating that the initial tonic "strives for" the concluding tonic. Such features include small-scale references—motives and non-structural progressions—that point to the concluding tonic prior to its structural establishment.[118] The relationship between the small-scale D and large-scale D in *Voiles* (Example 4e) is precisely analogous with such features. The "striving" to enlarge the harmony

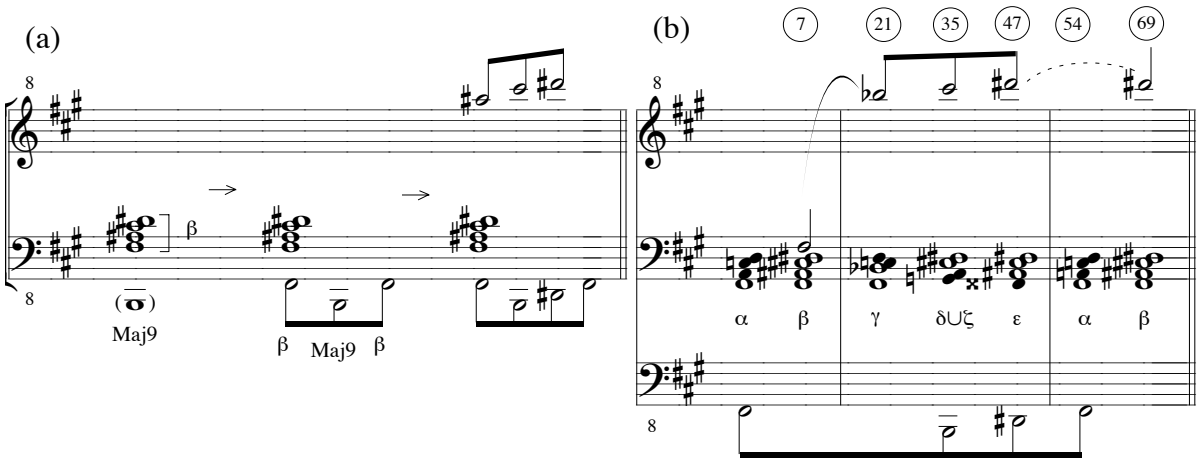
[117] **II** illustrates the existence of "related practices." The 10–2 (minor seventh – major ninth) arpeggiation occurs frequently in Debussy, for example. (This practice is adumbrated in conventional tonality by the elaboration of 9/7 chords.)

[118] See, for example, Carl Schachter's (1999: 260–88) analysis of Chopin's Fantasy op. 49. (I have also discussed this topic in Väisälä 2002: section 5.)

with fb-interval 4 is evident prior to its structural establishment, enhancing the pertinence of the referential harmony for the musical experience prior to its structural establishment.[119]

As a final example of referential harmonies formed through enlargement, I shall consider Debussy's *Ce qu'a vu le vent d'ouest*. With respect to both questions under discussion—the confirmation of the arpeggiating function and concomitant temporal effects—it is a more radical example than the preceding ones. Example 23 (reproduced from **III**: Example 9b–c) depicts the large-scale organization. It consist of two important structural aspects: the outer-voice structure, which projects the major ninth chord on B, and a chord progression based on semitonal neighboring relationships with "chord β" (F♯ major triad with added sixth). These aspects are coordinated by the fact that chord β is equal to the upper voices of the major ninth chord.

EXAMPLE 23. The large-scale structure of Debussy's *Ce qu'a vu le vent d'ouest*



The major ninth chord deviates from the referential harmonies of the preceding examples in that it *never* materializes at the surface. When the bass B appears in m. 35, upper tones do not form chord β which would complete the major ninth chord (instead, they form "δ∪ζ"). Moreover, the appearance of B is not adumbrated in small-scale events in a way comparable to the upper-voice D in *Voiles*. Is there then any kind of substantiation for viewing the bass line (F♯–B–D♯–F♯) in terms of an arpeggiation that enlarges the harmony, or is this reading as misleading as, for example, that in Example 21a? Fortunately, the state of affairs is not that bad. While the major ninth chord does not materialize in this *Prélude*, it occurs, in various structural functions and levels, in other pieces by Debussy (see **III**: Example 10). One foreground

[119] Similar considerations may also be applied to pieces based on several local referential harmonies. In the opening section of *Vers la flamme*, for example, there are clear non-structural references to C♯ and B, the members of the primary T_0A^{+9} outside the opening T_0U. The relationship between the C♯4–D4 and D4–C♯4 figures at the opening and in m. 27 "concretizes" the large-scale voice leading from D4 to C♯4. The bass motions from B♭ to B♮ in mm. 19 and 23 palpably "strive for" the corresponding large-scale motion prior to its structural establishment in m. 27. These relationships are illustrated in **II**: Example 8.

instance, from *La fille aux cheveux de lin*, is cited in Example 18 above. This passage also includes a foreground version of the B–D♯–F♯ (C♭–E♭–G♭) bass arpeggiation that occurs in the large-scale structure in *Ce qu'a vu le vent d'ouest*. The connections between these two consecutive *Préludes*—whose outward characters are utterly contrasting—are so striking that it seems unlikely that they are products of coincidence.[120]

All in all, if we extend our contextual considerations beyond *Ce qu'a vu le vent d'ouest*, we find support for viewing the major ninth chord as a harmony with a strong foothold in Debussy's musical imagination. Further support is given by considerations on rootedness, and on the general significance of rootedness for Debussy (**III**)—especially if we allow for the secondary root-supporting status of the major seventh (fb-interval 11) in the major ninth chord (see section 5.2.4 above). Unlike the "Martian" interpretations in Example 21, the reading in Example 23, based on the major ninth chord, has considerable explanatory power. It allows the structure to be related with several other pieces and passages in Debussy, including, in particular, the set of *Préludes*. The choice of the bass B is explained by the way it completes a pattern which has general importance in Debussy's music and which is also favored by root-based considerations. Moreover, no better way to explain the structural function of the bass B is available (at least the present author cannot think of one). It occurs concurrently with large-scale neighbors G and A, but cannot be described as "supporting" them in terms of any kind of consistent consonance–dissonance relationships.

I began this discussion of referential harmonies by identifying their double function as normative stable surface harmonies and as bases of the linear overall structure. In *Ce qu'a vu le vent d'ouest*, the harmonies fulfilling these two functions are not strictly identical: the former function is fulfilled by chord β and the latter by its enlargement, the major ninth chord. However, the tendency towards such enlargement may be understood as being inherent in β on the basis of Debussy's harmonic practices and root-based considerations. For a better illumination of the organization, it is best not to identify the referential harmony as either β or the major ninth chord. Rather, we need a conception of the referential harmony that allows both for the primary position of β and for its realized tendency to be complemented by the bass B to form the major ninth chord. The tones of chord β might be called "permanent" members of the referential harmony, whereas the bass B is an "associate" member. While *Ce qu'a vu le vent d'ouest* is the most radical example of the differentiation between "permanent" and "associate" members, a similar tendency is also evident in the preceding examples and may be generally relevant to considerations on post-tonal prolongation.

---

[120] In Debussy's preliminary sketch, the key of *La fille* is E♭ major instead of the final G♭. It is possible that changing the key had something to do with the desire to enhance the connections between the *Préludes* (although there may also be pianistic reasons for favoring the black-key tetrachord E♭–G♭–B♭–D♭).

6.2. NORMS RELEVANT TO THE FOUR CONDITIONS: SUMMARY AND ADDITIONAL REMARKS

The ways in which Straus's four conditions are met in the present studies are largely evident from the discussion in the preceding sections. However, a brief summary, which also contains some additional remarks, may be helpful for the reader.

Condition #1. Generally speaking, the norm of consonance is based on the similarity in construction with the referential harmony. How this similarity should be understood varies from case to case. Such variation concerns, for example, the extent and kind of permissible changes in registration (cf. section 4.1).

Similarity of construction does not necessarily mean sameness. In general, subsets or incomplete forms of the referential harmony are consonant, even though not all such forms are equally viable. Such viability depends on the importance of different intervals in the prevalent harmonic conception. For harmonic languages focused on bass-related intervals (FB; see section 4.1.2) the presence of the bass is more crucial than that of the upper voices. If, on the other hand, the focus is more on the adjacent intervals within a harmony (PCINT), the presence of inner voices may also be essential.

Apart from subset relationships, there are other kinds of similarities relevant to consonance, as exemplified by Schoenberg's 19/2 and Debussy's *Ce qu'a vu le vent d'ouest*. In the Schoenberg, there are two kinds of consonant chords, sharing features such as the systematic use of ro-interval 11 between registral layers, and the avoidance of ro-interval 1 (Example 13; **I**: 240–42). In the Debussy, local consonances are formed from the referential β chord (referential with the qualifications discussed above) by semitone shifts. Neither case is based on set-theoretical similarity. In both cases the referential harmony retains structural priority; other chords showing the discussed similarities with the referential harmony are consonances of a secondary rank (cf. section 3.2.3.1).

Condition #2. "Scale degree" systems emerge from the positions that the transpositions of the referential harmony assume in the contrapuntal large-scale embellishment of the referential harmony. This issue is extensively discussed in section 3.2.3.5 and requires no further comments. It should be repeated, however, that overall structures are not based on such "scale degree" systems in all of the present examples. The roles of chord construction and chord transposition on different scales of organization may be switched over, as discussed in section 3.2.3.1. In *Ce qu'a vu le vent d'ouest*, the large-scale organization is based on semitone-related chords shown by Greek characters in Example 23. They relate to the primary β by varying its construction rather than by transposition.[121] There is a large-scale arpeggiation of the

[121] Transpositional relationships occur only between the secondary harmonies: α and ε are transpositions of each other as are γ and δ.

referential harmony, including its "associate" member, the bass B, but this does not relate with any "scale degree" system based on transpositional relationships.

Condition #3. The embellishment types in the present analyses are similar to the conventional ones: arpeggiations, passing tones, complete and incomplete neighbors, and suspensions (which may be technically defined as incomplete neighbors with preparation). Regarding the special nature of bass-line embellishments in chord progressions, the reader is referred to section 3.2.3.5.

It may be worth noting that incomplete neighbors occur frequently in some of the examples. While incomplete-neighbor figures are in themselves "weaker" embellishments than complete neighbors and passing tones, they may be strengthened by a regular position in motivic repetitions (cf. section 3.2.3.3). For discussion of such motivic regularity, see especially the Berg and Webern analyses in **II**. Such motivic regularity may be understood as inducing work-specific specification for the embellishment condition.

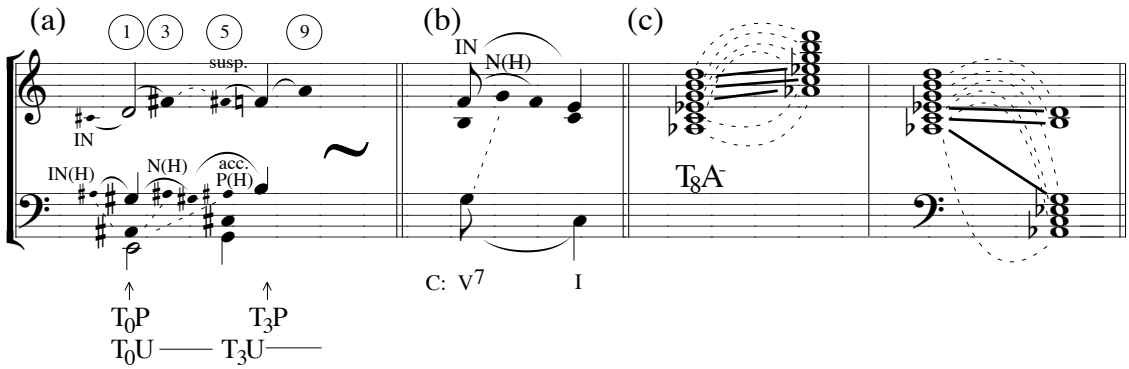
II (section 1.3) also identifies and discusses a type of incomplete neighbor which recurs in several compositions. This embellishment tone, called a *dimming tone*, is an incomplete neighbor that occurs after the main tone and lies a semitone lower. In the present text, examples of the dimming tone can be found in Examples 4c and 10b (in the former, it is indicated as IN(D), the notation generally used in **II**). Example 10b shows how the motivically important figure E–E♭–C, a descending arpeggiation elaborated by a dimming tone, occurs at two structural levels in Berg's op. 2/2. Example 4c shows the dimming-tone figure D♮–D♭ in Debussy's *Voiles*—a relationship with an extremely powerful expressive effect. Characteristically, the dimming tone functions as articulating the endpoint of the temporal presence of the main tone while prolonging it in a more abstract sense. It also has the function of adding sonorous variety. In both the Berg and Debussy examples, it involves a motion from ro-interval 4 to 3. In *Voiles* these intervals are formed in relation to the bass; in the Berg, they are formed in relation to an inner voice. Owing to such relationships, the sonorous effect of the dimming tone may be compared with tonal mixture. In the Berg and Debussy examples, the dimming tone also adds variety by effecting a digression from the governing whole-tone set.

Condition #4. In all examples, the harmony/voice-leading condition is fulfilled on the basis of the proximity principle of voice leading. Voice-leading intervals are small ("steps"), whereas arpeggiations are large ("leaps"). The "default" borderline between these categories goes between 2 and 3 semitones. Regarding possible octave generalizations and modifications of this borderline, see section 3.2.3.4. On the tendency to use larger intervals in non-arpeggiating functions in bass lines, see section 3.2.3.5.

One issue significant for embellishments and voice leading (conditions #3 and #4) remains to be mentioned: motion that concurrently involves "stepwise" and octave relationships. Since all the present referential harmonies include interval class 2 and some of them also include 1, "stepwise" embellishments within such harmonies may form octave (and

sometimes even unison) relationships with chord tones. In **II**, such embellishments are called h-neighbors and h-passing tones. In graphic notation, they are indicated by "(H)"; the octave relationships are shown by dotted lines. Example 24a illustrates these symbols for the opening of *Vers la flamme* (in Example 22a(v) these symbols are omitted).

EXAMPLE 24. Combinations of "stepwise" and octave relationships



The concept of h-neighbors and h-passing tones involves something of a paradox, because in one sense such tones are non-harmonic but in another sense they belong to harmony (cf. **II**: section 1.3). The A♯ that occurs in such figures at the opening of *Vers la flamme* represents fb-interval 6, which belongs to the opening harmony T_0U. One might argue that the A♯3 in the h-embellishments does not actually belong to the harmony because it does not occur in the same register as the A♯2 that represents fb-interval 6 the harmony. Such an argument is based on the idea that while chord U is defined in terms of fb-intervals, its actual realizations consist of pitches with definite registral location. However, this argument does not apply to all cases: there are some instances in which h-neighbors involve unison relationships instead of octaves. While the registral separation between the non-harmonic A♯3s and the harmonic A♯2 helps to clarify their functional distinction, a more definitive factor, and one with more general significance, is the clarity of the embellishment figures, which guarantees that the function of the embellishing A♯3s is determined on the basis of their horizontal, "stepwise" relationships more or less irrespective of the concurrent vertical octave relationships.

In most cases, the octave (unison) relationships that connect with h-embellishments are of little structural significance. That the paradox of "non-harmonic tones belonging to harmony" does not usually cause any practical problems for analysis is also illustrated by examples of conventional tonality, in which *seventh chords* are embellished by "h-neighbors"; see Example 24b. Such examples are commonplace in tonal music, even though one may regard them as being more radical than the present examples (such as 24a), owing to the violation of the consonance–dissonance condition in the structural order of tones.

While in cases such as Example 24a, octave relationships are structurally insignificant, there are also significant octave relationships, i.e., registral transfers, which may occur concurrently with "stepwise" embellishments. An especially striking example is to be found in *Voiles* (Example 4c). In the middle phases of the piece, the middle-register D has a double function as a registral transfer of the structural D7 and as a neighbor of E. This double function links with the special role of D in the temporal organization, the unconventional features of which are elaborated in the analysis in **II**.

Finally, there are cases in which octave relationships clearly take precedence in determining structural relationships. Example 24c reproduces Example 13 of **I**. The notation is different from the preceding examples: octave relationships are indicated by dotted slurs and semitone relationships by lines. The primary feature in this example is the registral transfer of chord "T_8A-" in its entirety.[122] Semitone motions have secondary significance in connecting different registral layers in the octave-related chords. Such a concurrent existence of two aspects of voice-leading is a feature that deviates from conventional tonality but does not jeopardize the clarity of structural relationships (**I**: 247).

6.3 ANALYTICAL RESULTS

While the functional norms related to the four conditions are different from conventional tonality, in other essential respects the analytical methodology follows customary Schenkerian practices. The structural relationships are indicated in voice-leading graphs by symbols familiar from Schenkerian analyses. Deviations from customary practices (most evident in **I**) are either explained, or—it is hoped—self-explanatory.

I will not reproduce the analytical results in **I–III** here but only point out some very general conclusions. Altogether, the analytical examples show, first, that post-triadic harmonies are capable of being prolonged in a more or less strict Strausian sense, second, that such possibilities have been realized in widely variable ways by some of the most significant 20th-century composers, and, third, that the theoretical principles discussed in this dissertation are crucial for revealing such prolongational organization.

The way in which the present analyses illuminate music is largely similar to that of conventional Schenkerian analyses. The prolongational approach is able to reveal aspects of "organic coherence" in a way not attainable by other methods. However, the revelation of prolongational structures is not only an end-in-itself. Such structures also serve as frames of reference for considerations on various other aspects, including central issues of musical

[122] These octave transfers do not occur as such in Schoenberg's op. 19/2, but the motion from T_8A to B (Example 10a) combines the downward transfer with whole-tone motions A♭–F♯ and G–F.

expression (cf. **II**: Introduction). For reasons of space, however, such considerations are made only occasionally in **I**–**III**.

An important question is, of course, the extent to which the present kind of principles are applicable to other post-tonal music (see discussions in **I**: section 5, **II**: section 3, and **III**). Actually there are two parts to this question. First, how relevant are the present principles for exploring whether prolongational organization occurs in post-tonal music? Second, to what extent will such exploration yield positive results? Regarding the first question, I would suggest that at least some of the present principles—such as the register-sensitive approach to harmony and voice leading—are generally significant (section 4.1.1). The second question, however, can be considered only in a most cautious way without extensive additional analytical studies. In the present studies, Debussy is the composer whose music is most extensively explored. These explorations speak for the utility of the present approach in the analysis of Debussy. However, the principles of post-tonal prolongation are by no means always realized with such purity as in *Voiles* and *Ce qu'a vu le vent d'ouest*, but are often combined with elements of more conventional (triadic) tonality (**III**). Scriabin is another composer to whose music the present kind of approach surely has wider applicability. *Vers la flamme* is in many respects characteristic of his late output.

In this connection, one may note that while the works analyzed in the present studies are *post-tonal* in the sense of not being governed by the triad, none of them is emphatically *anti-tonal* in the sense of avoiding any reminiscence of "tonal" elements. For example, it may be noted that in four cases—Schoenberg's op. 19/2, Scriabin's *Vers la flamme* (overall structure), Webern's op. 3/3, and Debussy's *Ce qu'a vu le vent d'ouest*—the referential harmony contains the major triad. On the other hand, in Berg's op. 2/2 there are associations between the referential harmony and a conventional dominant (**II**). One may ask whether such "non-antitonality" is generally characteristic of music to which the present principles apply.

In considering this question, it should first be noted that the associations of certain elements with the elements of conventional tonality does not play a structurally determinative role in the present analyses. In this respect the present approach sharply deviates from studies such as Baker 1990 and Cinnamon 1993 (see **I**: 246). For example, in Berg's op. 2/2—in which the allusions to conventional tonality are most explicit—the "tonic" alluded to is far from the structural background under the present interpretation (in contrast to Ayrey 1982).

While such explicit references to tonal music in a post-tonal context are thus of minor significance for the present analyses, some common features between the present repertoire and conventional tonality may arise from underlying principles that derive from psychoacoustics (**I**: 246). Under the proximity principle of spacing, "thirds" are the smallest possible harmonic intervals. The use of "tertian" harmonies (such as chord A in Schoenberg's op. 19; see Example 13) may be explained as a realization of such a basic possibility with no necessary reference to conventional tonality. Moreover, if rootedness has a role to play in the formation of

the referential harmony, this may naturally lead to the presence of the major triad in the harmony—even if some of the examples in **II** avoid this by the omission of the fifth. However, neither the proximity principle of spacing nor rootedness is indispensable for prolongation. Moreover, neither principle by any means necessarily leads to harmonies of tertian construction (as is the case in several examples in **II**).

It is thus possible that post-tonal prolongation may also be evident in music utilizing more radically atonal or anti-tonal harmonic vocabulary. On the other hand, it is also possible that the notion of prolongational hierarchies based on a primary harmony is not concordant with the aesthetic strivings of some trends of 20th-century music (cf. **II**: Section 3). Nothing definitive can be said on the basis of casual considerations of musical surface. Some principles—such as the distinction between "permanent" and "associate" members of harmony in *Ce qu'a vu le vent d'ouest* (section 6.1.2) or the supplementary principles of harmony in Webern's op. 3/1 (**II**)—that have been only tentatively introduced in the present studies might by further elaboration help to enlarge the scope of prolongational analysis to cases that at first sight seem resistant to it.

Even if the body of post-tonal compositions with as clear and pure prolongational overall structures as in the present examples proves to be limited, principles of the present approach may have utility in the description of compositions manifesting less "pure" organizational principles. Such utility may be evident in two areas. The first is music in which the overall syntax is not "emancipated" from conventional triadic norms but in which post-tonal formations assume temporary significance (see especially the closing discussion of Debussy's *L'Isle joyeuse* in **III**; see also section 5.2.3 above). The second area is post-tonal or atonal music, in which there are unlimited possibilities for norms relevant to the four conditions to emerge temporarily and partially even when not pervading the overall organization so as to sustain all-encompassing prolongational structures.

REFERENCES

Agmon, Eytan. 1990. “Music Theory As Cognitive Science: Some Conceptual and Methodological Issues.” *Music Perception* 7: 285–308.

Aldwell, Edward and Carl Schachter. 1989. *Harmony and Voice Leading*. Fort Worth: Harcourt Brace Jovanovich, 2nd ed.

Ayrey, Craig. 1982. “Berg's 'Scheideweg': Analytical Issues in Op. 2/ii.” *Music Analysis* 1: 189–202.

Baker, James. 1983. “Schenkerian Analysis and Post-Tonal Music.” In *Aspects of Schenkerian Theory*, ed. David Beach, 153–88. New Haven: Yale University Press.

———. 1986. *The Music of Alexander Scriabin*. New Haven: Yale University Press.

———. 1990. “Voice-Leading in Post-Tonal Music: Suggestions for Extending Schenker's Theory,” *Music Analysis* 9: 177–200.

Bregman, Albert S. 1990. *Auditory Scene Analysis*. Cambridge, Mass.: The MIT Press.

Burkhart, Charles. 1978. “Schenker's 'Motivic Parallelisms'.” *Journal of Music Theory* 22: 145–75.

Castrén, Marcus. 1997. “Joukkoluokitukseen perustuva sointuluokitus: perusperiaatteet ja esimerkkejä sovellusmahdollisuuksista.” *Sävellys ja musiikinteoria* 1997/1: 6–25.

Cinnamon, Howard. 1986. “Tonic Arpeggiation and Successive Equal Third Relations as Elements of Tonal Evolution in the Music of Franz Liszt.” *Music Theory Spectrum* 8: 1–24.

——— 1993. “Tonal Elements and Unfolding Nontriadic Harmonies in the Second of Schoenberg's Drei Klavierstücke, Op. 11.” *Theory and Practice* 18: 127–70.

Cook, Nicholas. 1987. “The Perception of Large-Scale Tonal Closure.” *Music Perception* 5: 197–206.

———. 1989. “Music Theory and 'Good Comparison': a Viennese Perspective.” *Journal of Music Theory* 33: 117–41.

Dibben, Nicola. 1999. “The Perception of Structural Stability in Atonal Music: The Influence of Salience, Stability, Horizontal Motion, Pitch Commonality, and Dissonance.” *Music Perception* 16: 265–94.

Friedmann, Michael L. 1982. “Approaching Debussy's 'Ondine'“ *Cahiers Debussy* 6: 22–35.

Gjerdingen, Robert. 1999. “An Experimental Music Theory.” In *Rethinking Music*, eds. Nicholas Cook and Mark Everist, 161–170. Oxford: Oxford University Press.

Hartmann, William M. 1997. *Signals, Sound, and Sensation*. Woodbury: AIP Press.

Houtsma, A. J. M. and J. L. Goldstein. 1972. "The Central Origin of the Pitch of Complex Tones: Evidence from Musical Interval Recognition." *The Journal of the Acoustical Society of America* 51: 520–29.

Houtsma A. J. M. and J. Smurzynski. 1990. "Pitch Identification and Discrimination for Complex Tones with Many Harmonics." *The Journal of the Acoustical Society of America* 87: 304–10.

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

Katz, Adele T. 1972. *Challenge to Musical Tradition*. New York: Da Capo Press. First published 1945.

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

Laufer, Edward. 1981. "Review of *Free Composition*." *Music Theory Spectrum* 3: 158–84.

Laufer, Edward. 1991. *Schoenberg's Klavierstück, Op. 33a: A Linear approach*. (A conference paper presented to a joint meeting of Arnold Schoenberg Institute/Music Theory Society of New York State, October 4, 1991.)

Lerdahl, Fred. 1989. "Atonal Prolongational Structure." *Contemporary Music Review* 4: 65–87.

———. 1997. "Issues in Prolongational Theory: A Response to Larson," *Journal of Music Theory* 41: 141–55.

———. 1999. "Spatial and Psychoacoustic Factors in Atonal Prolongation," *Current Musicology* 63: 7–26.

———. 2001. *Tonal Pitch Space*. New York: Oxford University Press.

Lerdahl, Fred and Ray Jackendoff. 1983. *A Generative Theory of Tonal Music*. Massachusetts: The MIT Press.

Lester, Joel. 1970. "A Theory of Atonal Prolongations as Used in an Analysis of the Serenade, Op. 24 by Arnold Schoenberg." Diss., Princeton University.

Lewin, David. 1986. "Music Theory, Phenomenology, and Modes of Perception," *Music Perception* 3: 327–92.

Moore, Brian C. J., Brian R. Glasberg, and Robert W. Peters. 1985. "Relative Dominance of Individual Partials in Determining the Pitch of Complex Tones." *The Journal of the Acoustical Society of America* 77: 1853–60.

Morgan, Robert P. 1970. "Dissonant Prolongation: Theoretical and Compositional Precedents." *Journal of Music Theory* 20: 46–91.

Morris, Robert. 1995. "Equivalence and Similarity in Pitch and their Interaction with Pcset Theory." *Journal of Music Theory* 39: 207–44.

Morrison, Charles. “Prolongation in the Final Movements of Bartók's String Quartet No. 4.” *Music Theory Spectrum* 13: 179–96.

Oster, Ernst. 1960. “Re: A New Concept of Tonality (?).” *Journal of Music Theory* 4: 85-98.

Parncutt, Richard. 1988. “Revision of Terhardt's Psychoacoustical Model of the Root(s) of a Musical Chord.” *Music Perception* 6: 65–94.

———. 1996. “Praxis, Lehre, Wahrnehmung. Kritische Bemerkungen zu Roland Eberlein: 'Die Entstehung der tonalen Klangsyntax'.” *Musiktheorie* 11: 67-79.

Pearsall, Edward. 1991. “Harmonic Progressions and Prolongationa in Post-Tonal Music,”*Music Analysis* 10: 345–56.

Perle, George. 1984. “Scriabin’s Self-Analysis,” *Music Analysis* 3 (1984): 101-22.

Persichetti, Vincent. 1962. *Twentieth-Century Harmony: Creative Aspects and Practice*. London: Faber and Faber.

Plomp, R. and W. J. M. Levelt. 1965. “Tonal Consonance and Critical Bandwidth.” *The Journal of the Acoustical Society of America* 38: 548–60.

Rothgeb, John. 1978. A review of *Beyond Schenkerism: The Need for Alternatives in Music Analysis* by Eugene Narmour. *Theory and Practice* 3: 28–42.

Rothstein, William. 1991. “On Implied Tones.” *Music Analysis* 10: 289–328.

Salzer, Felix. 1982. *Structural Hearing: Tonal Coherence in Music*. New York: Dover Publications. First published 1952.

Schachter, Carl. 1999. *Unfoldings: Essays in Schenkerian Theory and Analysis*. Ed. Joseph N. Straus. New York: Oxford University Press.

Schenker, Heinrich. 1954. *Harmony*. Ed. Oswald Jonas. Trans. Elisabeth Mann Borgese. Chicago: The University of Chicago Press. German original 1906.

———. 1996. *The Masterwork in Music: A Yearbook: Volume 2*. Ed. William Drabkin. Trans. Ian Bent, William Drabkin, John Rothgeb, and Hedi Siegel. Cambridge: Cambridge University Press. German original 1926.

———. 1979. *Free Composition*. Trans. Ernst Oster. New York: Schirmer Books. German original 1935.

Schoenberg, Arnold. 1978: *Theory of Harmony*. Trans. Carter, Roy E. London: Faber Music. Translation based on the third German edition 1922. German original 1911.

Snyder, John L. 1991. “Schenker and the First-Movement of Mozart’s Sonata, K. 545: *An Uniterrupted Sonata-Form Movement?*.” *Theory and Practice* 16: 51–78.

Straus, Joseph. 1987. “The Problem of Prolongation in Post-Tonal Music.” *Journal of Music Theory* 31: 1–22.

———. 1997a. “Voice Leading in Atonal Music.” In *Music Theory in Concept and Practice*, ed. James M. Baker, David W. Beach, and Jonathan W. Bernard, 237–74. Rochester: University of Rochester Press.

———. 1997b. “A Response to Larson.” *Journal of Music Theory* 41: 137–40.

———. 2003. "Uniformity, Balance, and Smoothness in Atonal Voice Leading." *Music Theory Spectrum* 25: 305–52.

Temperley, David. 2001a. *The Cognition of Basic Musical Structures*. Cambridge, Mass.: The MIT Press.

———. 2001b. "The Question of Purpose in Music Theory: Description, Suggestion, and Explanation." *Current Musicology* 66: 66–85.

Terhardt, Ernst. 1974. "Pitch, Consonance, and Harmony." *The Journal of the Acoustical Society of America* 55: 1061–69.

———. 1984. "The Concept of Musical Consonance: A Link Between Music and Psychoacoustics." *Music Perception* 1: 276–95.

Travis, Roy. 1959. "Toward a New Concept of Tonality." *Journal of Music Theory* 3: 257–84.

———. 1966. "Directed Motion in Schoenberg and Webern." *Perspectives of New Music* 4: 84–89.

Travis, Roy. 1990. Communication. *Journal of Music Theory* 34: 380.

Väisälä, Olli. 2002. "Review of Timothy L. Jackson and Veijo Murtomäki (ed.), *Sibelius Studies*." *Music Theory Online* 8, No. 4.

Wagner, Naftali. 1990. "Investigations in Tonal Structural Perception (A Preliminary Study)." *Israeli Studies in Musicology* 5: 147–70.

Wilson, Paul. 1992. *The Music of Béla Bartók*.." New haven and London: Yale University Press.

Winkler, I., M. Tervaniemi, and R. Näätänen. 1997. "Two separate codes for missing-fundamental pitch in the human auditory cortex." *The Journal of the Acoustical Society of America*. 102: 1072–82.

Zenck-Maurer, Claudia. 1974. *Versuch über die Wahre Art, Debussy zu analysieren*. München–Salzburg: Musikverlag Emil Katzbichler.

In the printed dissertation, photocopies of articles **I–II** and a manuscript of **III** follow as attachments.

I Väisälä, Olli. 1999. "Concepts of Harmony and Prolongation in Schoenberg's Op. 19/2." *Music Theory Spectrum* 21: 230–59.

II Väisälä, Olli. 2002. "Prolongation of Harmonies Related to the Harmonic Series in Early Post-Tonal Music." *Journal of Music Theory* 46: 207–83.

III Väisälä, Olli. 2006. "New Theories and Fantasies on the Music of Debussy: Post-Triadic Prolongation in *Ce qu'a vu le vent d'ouest* and Other Examples." In *Selected Essays from the Third International Schenker Symposium*. Hildesheim: Georg Olms, 2006.