Alan Dodson[*]

Performance and Hypermetric Transformation: An Extension of the Lerdahl-Jackendoff Theory[*]

To listen to the audio files, you will need the **free** RealAudio player.

Get the **free** DjVu plug-in to best view the examples.

ABSTRACT: In the course of the introductory commentary on hypermeter in *A Generative Theory of Tonal Music* (*GTTM*), Lerdahl and Jackendoff discuss the opening measures of Mozart's Symphony No. 40 in G Minor, a hypermetrically ambiguous passage in which "the performer's choice . . . can tip the balance one way or the other for the listener." Through reflections on concepts from more recent psychological inquiry into performance, and on the interpretations of the passage that are projected in four well-known recordings of the Symphony, I will develop a set of theoretical principles that describe the "balance-tipping" effects of performance-specific elements on hypermetric structures inferred by the listener. This special case will lead to a more general reconsideration of the place of performance in the design of the Lerdahl-Jackendoff theory. The article proceeds in five parts: (1) an introduction to the main theoretical concepts to be discussed, including a brief consideration of current debates with which the study intersects; (2) a critical discussion of the relationship between hypermeter and performance that is proposed in *GTTM*; (3) an attempt at extending the theory of accent types to include a special class of phenomenal accents that is under the performer's control; (4) a close reading of four recordings, facilitated by quantitative performance analyses, and an attempt at explaining their hypermetric patterns as transformations of perfectly regular underlying structures; and (5) concluding remarks of a more general nature on the relationship between structure and performance in *GTTM*.

References

---

[1] Introduction

[1.1] During the past few decades, theorists have been rethinking the concept of musical meter at a fundamental level. In recent literature, meter has been described not as a static pattern of accents intrinsic to the musical object, but rather as a dynamic process whose existence depends on human participation.[(1)] This shift can be demonstrated through the following contrasting definitions of meter, drawn from the first and second editions of *The New Grove*, respectively:

1980 ed.: The organization of the notes in a composition, or a section thereof, with respect to time, in such a way that a regular pulse made up of beats can be perceived and the time span occupied by each note can be measured in terms of these beats; in addition, the beats are grouped into larger units called bars, within which the number of beats is always the same.[(2)]

2001 ed.: [T]he temporal hierarchy of subdivisions, beats and bars that is maintained by performers and inferred by listeners which functions as a dynamic temporal framework for the production and comprehension of musical durations. In this sense, meter is more an aspect of the behaviour of performers and listeners than an aspect of the music itself.[(3)]

Both of these definitions allude to music perception, and the distinction between them might be understood to reflect a shift in the orientation of music perception research that occurred in the 1980s. Traditionally, perception was understood as the relatively passive mental reproduction of a given physical stimulus, but it is now generally considered to be an active, constructive (as opposed to reconstructive) process.[(4)] The revised definition of meter might also be understood against the backdrop of a similar, though perhaps more gradual, paradigm shift in music theory and in music scholarship generally. Largely in response to far-reaching methodological and ideological critiques, notably Joseph Kerman's *Contemplating Music* and subsequent work by postmodernists and gender theorists, we seem to have become increasingly skeptical of the objectification of musical works and the quasi-taxonomic interpretation of their elements.[(5)] In this light, it seems especially appropriate that today's theorists should emphasize the subjective contingency, as opposed to the autonomy, of meter.

[1.2] In the present article, I will explore some of the theoretical and analytical ramifications of the recent reconceptualization of meter, especially the effects of performing nuance on meter perception. I will take as my point of departure the discussion of the interaction of meter

and performance in Fred Lerdahl and Ray Jackendoff's *A Generative Theory of Tonal Music* (*GTTM*).[6] I have chosen this text for several reasons. First, *GTTM* has proven to be among the most convincing sustained exercises in psychologically informed music theory.[7] Second, it has been highly influential among theorists interested in meter and among music psychologists interested in performance, so it can be seen as a well-established link between the literatures on which I will draw.[8] Third, the language of *GTTM* is extremely clear, and the theory's methodological foundations and orientation have been explained thoroughly.[9] For example, the authors introduce the theory as an attempt to capture in formal language "the musical intuitions of a *listener* who is experienced in a musical idiom," rather than aspects of the composer's intentions or intrinsic properties of musical works.[10] Furthermore, the authors are frank about the theory's shortcomings, thus reminding the reader that *GTTM* is incomplete and suggesting that it has room for expansion and refinement.[11] Fourth, the passage from *GTTM* that I will take as my point of departure is related to three topics that have remained controversial during the nearly two decades since the publication of *GTTM*. These include discussions on the interrelationship between musical structure and performance, on the epistemological foundations for comparing different performances of the same piece, and on the degree to which hypermeter can be irregular. (By "hypermeter," I mean the projection of a pattern of strong and weak beats across units larger than one measure.[12]) Of these three controversial topics, I will say most about the third, but I would like to offer a few remarks on the first two issues at the outset.

[1.3] Commentaries on the relationships between musical structure and performance form an increasingly prominent genre in the literature of music theory.[13] In a recent critique of the rhetoric most often employed in this genre, and in music analysis generally, Nicholas Cook draws attention to the dogmatic, prescriptive tone that is typical of analysts' suggestions to performers.[14] As an alternative, Cook recommends that we strive for a more balanced dialogue in which neither the analyst nor the performer is considered to have the upper hand. By invoking J. L. Austin's theory of speech-acts, Cook proposes that we might begin to read analyses not only as truth-claims (which Austin terms "constative utterances") but also as acts of persuasion ("performative utterances"), thus highlighting a deep similarity between musical performance and analytical writing.[15] Building on earlier polemics by Tim Howell and Lawrence Rosenwald, Cook also encourages analysts to listen to performances and recordings and to make descriptive, as opposed to prescriptive, remarks on the relationships between structure and performance.[16] In another recent essay, Joel Lester demonstrates this strategy convincingly by showing

parallels between various recordings and analyses in the case of structurally ambiguous passages. For example, Lester shows how two different analytic perspectives on a Mozart minuet correspond to details in recordings of the piece by Lili Kraus and Vladimir Horowitz. [(17)] I will adopt a similar discursive strategy in my readings of hypermeter in the opening of a Mozart symphony by showing structural models that seem to fit best with the nuances of four different recordings, but I will also go further and explore in purely theoretical terms how we might begin to account for the diversity encountered in a comparison of diverse performances.

[1.4] An important methodological question faces those of us who would like to theorize the interrelationships of performance and structure, and in particular the effects of different performances on the listener's perception of structure: what is the conceptual framework within which comparisons between performances can be made at all, that is, how can we account (in theoretical, as opposed to historical or stylistic, terms) for the differences that we encounter in comparative listening? At one extreme, scholars who aspire to positive knowledge, who contend that we should aim for a single, comprehensive analysis of any work, might be expected to evaluate the quality of a performance on the basis of the convincingness of the analysis that it seems to reflect and to consider the differences between performances to be a result of differences between performers' levels of insight. [(18)] At the other extreme, those who have embraced pluralism and liberalism would, on principle, question the validity of any theory that encourages positivistic thinking, and might instead attribute differences between interpretations to the diverse and unpredictable personal and cultural contexts of each performance. [(19)] My own approach lies somewhere between these two extreme positions, because I would like to bypass certain methodological obstacles latent in each: the former paradigm (which we might call the positivist "competition" model) is inconsistent with our ability to recognize two very different performances of the same work to be more-or-less equally convincing, and the latter (which we might call the pluralist "free-for-all" model) seems too broad to be commensurate with the project of discussing differences between performances primarily in terms of the perception of structure.

[1.5] My approach is based on a third paradigm, which I call the "alternative stabilization" model of performance comparison. This paradigm will allow me to theorize the differences between performances in structural terms, but to do so without privileging one performance over the others. At this point, I will describe its premises in only the most general and practical terms. When analyzing a work through *reading* it from the score, one sometimes encounters genuine ambiguities, passages in

which some aspect of the structure perplexingly eludes a single preferred interpretation. [20] In *listening* to performances and recordings, however, much more information is available than would be found in the score alone. As Nicholas Cook recently reminded us, scores are not only potential objects of analysis, but also function as "scripts" in many of the social contexts in which notation is sounded out--scripts to which a performance serves as a sort of supplement. [21] I am interested in situations in which this supplement removes, or at least reduces, the degree of ambiguity encountered in reading the score, such that the listener is unaware of the ambiguity, or at least less inclined than the reader to describe the structure in question as ambiguous. The performer must choose only one of the possible interpretations, so naturally performances can be compared to one another regarding the choices made by performers. [22] In the case of works or passages for which multiple versions exist in notation (e.g., "ossia" passages in Romantic piano music) or for which there is a pronounced improvisatory element (e.g., Corelli's slow movements, jazz), this stabilizing effect is perhaps most obvious. [23] In the case of hypermeter, however, it is not the notes that are chosen, but rather the subtle expressive nuances of a performance, including details of timing, dynamics, and timbre, that provide the stabilizing effect.

[1.6] Clearly, the phenomenon I am describing is incommensurate with the (Platonist) notion that for every work there is a singular ideal form. It would seem very difficult to reconcile the notion of alternative stabilization with a theory heavily informed by Western aesthetics and metaphysics. Heinrich Schenker, for example, insisted that for every masterwork there is a single correct interpretation of the voice-leading structure, and a rejection of this claim would amount to a fundamental change to the philosophy of Schenkerian theory. [24] Though I find the question of the interaction of musical ontology and performance fascinating, it is in fact tangential to this study, for I am thematizing perceptions, not speculative metaphysical fictions about the origins of a work or performance. [25] In my view, the issue of ambiguity and stability in *GTTM* should instead be discussed in relation to the theory's grounding in Gestalt psychology, a school of thought that sought to explain the coherence of things without recourse to idealism. [26]

[1.7] Contrary to popular belief, the leading exponents of Gestalt psychology did not base their research on the cliché, "the whole is greater than the sum of its parts." Rather, they claimed that the perception of the whole is categorically different from a summation of perceived parts. [27] Max Wertheimer was the first psychologist to argue that percepts are not put together piecemeal but rather are understood intuitively and immediately according to basic aesthetic principles, such that one should

proceed from the whole to the parts rather than the reverse in any attempt to theorize perception.[28] Wertheimer's most basic principle is the "Law of *Prägnanz* [precision]," which was most clearly expressed by his colleague Kurt Koffka: "Psychological organization will always be as 'good' as the prevailing conditions allow. In this definition the term 'good' is undefined. It embraces such properties as regularity, symmetry, simplicity, and others which we shall meet in the course of our discussion." Some of these "others" include "unity, uniformity, good continuation, simple shape, and closure," all of which are discussed at some length in Koffka's encyclopedic *Principles of Gestalt Psychology*.[29] Lerdahl and Jackendoff explain that the Law of *Prägnanz* is closely associated with their Preference Rules (PRs), whose main purpose in *GTTM* is to capture the sense of stability associated with the hierarchical organization of tonal music. (I will have more to say on PRs below.) In fact, they claim that "the preference rules in effect constitute an explicit statement of the Law of *Prägnanz* as it applies to musical perception."[30] The intricacies of perceptual stability and instability are perhaps best illustrated in the work of the Edgar Rubin and his circle, who are remembered for their studies on the perception of ambiguous visual stimuli.

[1.8] A few variations on Rubin's famous "faces/vase" drawing will serve as a first attempt at a visual analogy for what I have termed the alternative stabilization phenomenon (Example 1[31] [DjVu] [GIF]). Let us say we have a score in which two interpretations of the location of a phrase boundary seem equally tenable to the *reader*. This experience is a bit like looking at Rubin's drawing (Example 1a) and being unable to decide whether it depicts two faces or a vase. The location of the phrase boundary might be less ambiguous from the *listener's* perspective, however, because performing nuances, like a few embellishments to Rubin's drawing, can increase the perceptual stability of either interpretation (see Example 1b–c). Another visual analogy, which uses a subtler manipulation and also eliminates the representational component of the drawing, thus bringing the analogy closer to *GTTM* (a formalist theory), comes from the work of Rubin's student Paul Bahnsen, who examined the perceptual effects of the shape of the border between alternating black and white regions. Perception might be expected to be unstable when all borders are parallel or all borders are irregular, but it is highly stable when borders of either the black or white regions are symmetrical (as in Example 1d–e). This might be seen as an analogy for the projection of phrase boundaries through the use of "phrase rubato," which (perhaps coincidentally) is itself symmetrical, for the performer typically begins the phrase relatively slowly, accelerates to the climax, and decelerates at the end of the phrase.[32] I should add that some ambiguous figures, like some

musical passages, have more than two possible interpretations. For
instance, a single curved line (Example 1f) might be seen as the border
of a figure to the left, the border of a figure to the right, or simply
a line superimposed on a continuous background. This line is referred to
as a "tristable" figure, since it can be interpreted in three different
ways. Still more extravagant is the ambiguity found in some multistable
mosaic patterns (Example 1g). I will leave it to the reader to imagine
some manipulations that would stabilize patterns like Example 1(f) and
(g).

[1.9] We are now ready to explore some of the ways *GTTM* might be developed
to account more fully for the impact of performance on the listener's
experience of hypermeter. In Part 2, I will examine Lerdahl and
Jackendoff's position on this subject, which appears in the context of
an argument against the perception of irregular hypermeter. In Part 3,
I will attempt to revise the classification of accent types in *GTTM* to
include a category that I call "Phenomenal Micro-accents" (PMs) and
discuss the importance of these accents in the case of otherwise unstable
metric and hypermetric structures. In Part 4, I will present analyses of
hypermeter in the opening measures of Mozart's Symphony No. 40 in G Minor
in four different recordings and illustrate the stabilizing effects of
their different patterns of PMs, and I will also develop formal
transformation rules that model the listener's ability to hear the
irregularities in these patterns as modifications of an inferred,
perfectly regular metrical pattern. Finally, in Part 5, I will offer a
more general critique of Lerdahl and Jackendoff's position on the
relationship between performance and their analytic process, and I will
also identify some avenues for further research.

[2] Hypermeter and Performance in GTTM

[2.1] In *GTTM*, the term "Metrical Structure" (MS) refers to accent
patterns within the measure (sometimes called "bar meter") as well as
meter-like organization across spans smaller than measures (i.e.,
quasi-metric subdivisions of beats) and larger than measures (i.e.,
hypermeter).[33] In Lerdahl and Jackendoff's ingenious notational system
for MS, a dot under the score represents the point in time corresponding
to a beat, and the number of dots in a given column indicates the accentual
strength at that point relative to other beats in the hierarchy.[34] Their
analysis of the opening of a Beethoven scherzo (Example 2[35] [DjVu] [GIF])
demonstrates this system well. The local triple meter is shown by the
presence of a dot at level (b) under every third dot at level (a). Three
levels of hypermeter, all duple, can be inferred by comparing the dots
of the remaining pairs of adjacent levels in the example. Levels (b) and

(c) show five two-bar hypermeasures, (c) and (d) show two full four-bar hypermeasures, and levels (d) and (e) show a single full eight-bar hypermeasure. For simplicity, I will refer to hypermetric levels by number, with H_1 being the first level of hypermeter, H_2 the second, and so on.

[2.2] Each of the four main components of *GTTM* (of which MS is one) is governed by two sets of rules. The practice of developing rule systems stems from the system of transformational linguistics developed by Noam Chomsky, Jackendoff's mentor.[36] These rules are supposed to model the largely automatic cognitive processing by which musical surfaces are parsed. Many of the principles formalized in the Rule Index of *GTTM* come from Gestalt psychology and from more recent research in music perception. Preference Rules (PRs) are designed to capture the way experienced listeners interpret a unique combination of structural details.[37] As mentioned above, PRs are concerned mainly with clarifying perceptual stability, and they are applied to each piece *ad hoc* after the more general constraints captured by Well-Formedness Rules (WFRs) have been applied. Metrical Well-Formedness Rules (MWFRs) determine the number of levels and their patterns of strong and weak beats, while Metrical Preference Rules (MPRs) encourage the most logical alignment of this pattern with events at the musical surface such as dynamic accents, long notes, and cadences.

[2.3] Lerdahl and Jackendoff mention hypermetric ambiguity in their introductory remarks on MS in the opening measures of Mozart's Symphony No. 40 in G Minor, a passage they consider to be "not untypically complex" in its metrical organization.[38] They use this passage to demonstrate that hypermeter is often restricted to levels close to the foreground, in opposition to the view that hypermeter can operate at even the deepest structural levels.[39] Near the beginning of the passage, odd-numbered measures are more strongly accented than even-numbered measures, but the reverse is true by m. 20. The only way to rationalize this, according to the authors, is to infer a single three-bar hypermeasure in the midst of a series of two-bar hypermeasures. They offer two equally plausible readings of where this eccentric hypermeasure might be found (see Example 3[40] [DjVu] [GIF]). In reading A, the shift to even-measure accents occurs at what seems to be the latest possible moment, the downbeat of m. 14, while in reading B, the shift occurs at the first sign of change, the downbeat of m. 10.[41]

[2.4] Lerdahl and Jackendoff then claim that, despite the instability created by the irregular hypermeasure, "the performer's choice . . . can tip the balance one way or the other for the listener."[42] In other words, a performance might evoke a clear sense of hypermeter in the passage, but no such clarity is available on the basis of the score alone. (This claim calls to mind Example 1, which illustrates how a perceptually unstable

figure can be stabilized through subtle manipulations. See especially Example 1b-e.) Although the authors claim that hypermeter can exist in this passage, they abstain from working out in practical terms precisely *how* features of performance might be understood to stabilize the hypermeter. Instead, they decide only to consider regularly spaced beats in their theory of MS, at least at the *tactus* (most salient metrical level) and immediately larger levels (MWFR 4). They do propose a special "Metrical Deletion" rule that deals with irregularities arising from situations like phrase elision and overlapping, but neither the rule index nor the formal discussion of MS gives evidence that the analytic implications of the potential "balance-tipping" effect of performance have been accounted for.[43]

[2.5] Although the variability of performance provides Lerdahl and Jackendoff with a pragmatic objection to irregular hypermeter, and consequently to deep levels of hypermeter in most contexts, it is not clear that this difficulty amounts to a *theoretical* impasse. I would suggest that most experienced listenersapparently including Lerdahl and Jackendoff themselvesperceive some degree of quasi-metrical organization in the Mozart excerpt shown in Example 3 ([DjVu] [GIF]) across spans larger than one measure. In order that the theory might better capture the experienced listener's intuitions, I consider it worthwhile to develop in formal terms the "balance-tipping" effect that Lerdahl and Jackendoff mention. Clearly, irregular hypermeter is more complex than regular hypermeter, and thus more difficult to formalize, but it seems to me that this complexity ought to be confronted rather than avoided. In particular, I would argue that the degree of hypermetric irregularity typically encountered in performances of the Mozart excerpt is not so extreme as to obviate any sense of hypermeter; this excerpt seems to lie in the vast "grey area" between regular and random proportions, between the obvious and the incomprehensible.

[2.6] Gestalt psychology again provides a useful analogy. According to one interpretation of the Law of *Prägnanz*, the purpose of the act of perception is to simplify or regularize the information given in the stimulus if it is possible to do so (Example 4 [DjVu] [GIF]).[44] Thus, a perfectly regular shape such as the one shown in Example 4a may be altered to some degree, as in Example 4b-c, without causing the lack of integrity found in a random collection of line segments, such as Example 4d. This very principle enables Lerdahl and Jackendoff to explain away the subtle durational irregularities found at the beat-to-beat level in performance; they note that listeners are able to infer a regular pulse underlying a musical surface transformed through *tempo rubato*.[45] They do not, however, explore the functioning of this type of manipulation at the level of the

measure or hypermeasure, and the reason for this again seems to be a practical rather than a theoretical problem, namely, how to determine with precision what should be included in the inferred metrically regular musical surface that preceded the transformation.

[2.7] In the formal language of *GTTM*, the phenomenon I am describing could best be captured through a series of "transformational rules." This category of rules, which is separate from the WFRs and PRs, is loosely based on grammatical transformations such as the change from active to passive voice or from present to past tense.[46] The transformational rules in GTTM reverse the changes to the musical surface brought about by processes like phrase overlapping and elision.[47] In order to adhere to this notationally oriented practice, our transformational rules for hypermetric irregularities would need to be able to generate entire measures of inferred music. In an earlier attempt at applying Chomsky's principle of grammatical transformation to music, Leonard Bernstein did just that: he recomposed the opening of Mozart's Symphony No. 40 in such a way that regular metrical structures are found at levels H_1 and H_2, that is, hypermeter at the level of the double- and quadruple-measure (Example 5[48] [DjVu] [GIF] Audio[49] [Modem] [Broadband]). Bernstein's newly composed material (mm. 1–3, 15–16, 19–20, 23–24, 28–31, and 34–36) amounts to sixteen additional measures, making it two-thirds longer than the passage on which it is based. Lerdahl and Jackendoff claim that this type of approach is untenable, because it is too "hypothetical," by which I think they mean too far removed from the actual listening experience, and because it is so arbitrary in its details. While I agree with this assessment, I nevertheless find Bernstein's approach thought-provoking. As I mentioned near the beginning of this paper, meter is now regarded primarily as a construct created by the performer and listener, rather than an inherent and fully determined property of a musical work or score. One property of this construct appears to be that, once formed, it can be separated substantially from the musical surface. For instance, once we know a piece, we can imagine or even physically "feel" its metrical pattern without also imagining other parameters, such as the pitch materials. Indeed, psychologists have found empirical support for the human cognitive ability to form abstract, hierarchical mental representations of musical meter, and this finding seems highly compatible with Lerdahl and Jackendoff's MS theory and its associated analytic notation.[50] In my view, then, it should be sufficient to show the transformation of irregular hypermeter in an abstract sense rather than proposing transformations at the level of the musical surface. The schematic representation of MS in a given passage can easily be compared to a hypothetical, perfectly regular MS, and processes like metrical expansion and truncation can be reversed.[51] For instance, it is clear that

both of the MS interpretations of the Mozart example that Lerdahl and Jackendoff offer (see Example 3 [DjVu] [GIF]) contain a single three-bar hypermeasure amid a stream of two-bar hypermeasures. A transformational rule to regularize this pattern, thus modeling the simplifying function of the Law of *Prägnanz* in this context, would merely need to reverse this metrical expansion. (I will propose such a rule below; see paragraph 4.6.) The location of the departure from the regular pattern will depend on the performance, but we will nevertheless hear it as a departure *from something* that is more-or-less regular, stable, and comprehensible. That "something" is the underlying hypermeter.

[2.8] It is tempting to accuse Lerdahl and Jackendoff of succumbing to the influence of an insidious aesthetic bias, the so-called "autonomy ideology," on the grounds that they are reluctant to formalize the dependency of an analysis on a specific performance. That is, one could speculate that their "balance-tipping" analogy (see the passage from *GTTM* quoted in paragraph 2.4) is abandoned in their theoretical discussion because it is incommensurable with contemplating the inner workings of an autonomous (i.e., radically independent) musical work. This accusation would be unfair, however, given that Lerdahl and Jackendoff explicitly avoid the discussion of aesthetics; in subsequent publications, they have repeatedly pointed out that *GTTM* theorizes aspects of the comprehensibility of a work, not its aesthetic qualities, and that these two parameters are not always closely related.[52] *GTTM* might be regarded as a method for developing the simplest interpretation of a tonal work's hierarchical dimensions, for bringing the interpretation as close to the Gestalt ideal of "good form" as possible, such that a *GTTM*-style analysis can be thought of as representing structural intuitions that operate on the most perceptually "stable" events in a piece of tonal music. In other words, although it is a formalist theory, it attempts to escape idealism and instead to account for musical intuitions in terms of psychological principles drawn mainly from the work of the Gestalt school.[53] In this light, it would seem most appropriate to object to the abandonment of the balance-tipping model on the grounds of the Law of *Prägnanz* (see paragraph 1.7), rather than an aesthetic bias.[54] As we have seen, this law begins, "Psychological organization will always be as 'good' as the prevailing conditions allow." I would argue that the limiting factorthe "prevailing conditions" alluded to in this definitionshould at least potentially include all of the information contained in the aural stimulus, including features specific to an individual performance. The main obstacle to developing Lerdahl and Jackendoff's balance-tipping model would therefore seem to be not an ideological conflict but rather the practical difficulty of describing the elements of a performance with adequate precision.

[2.9] It has long been recognized that the projection of a structural interpretation is one function of the expressive nuances that performers add in their realizations of scores. These nuances are sometimes termed "systematic variations" (abbreviated SYVARs), because they can be described in quantitative terms as patterns of departures from mechanical regularity in a given domain, such as speed or loudness.[55] In the case of an unambiguous structure, it would be relatively unproblematic to assume a direct, linear connection between the score and the "musical surface" (i.e., the aural presentation of the music), so long as the hypothetical performer adheres to the same SYVAR conventions that the listener has absorbed through aural experience with the musical idiom (see Example 6a [DjVu] [GIF]). In such cases, the absence of PRs pertaining to performing nuances in the rule index of *GTTM* seems unproblematic; in a conventional performance, structural interpretations are projected rather than obfuscated or contradicted.[56] In the case of an ambiguous structure, however, the musical surface that is presented to the listener is often much clearer than the score itself.[57] It should follow that, in the case of ambiguous structures, the role of the performer's interpretive preference, as expressed through SYVARs, deserves some consideration (see Example 6b-c). The fact that these SYVARs cannot be predicted definitively on the basis of a score makes them no less relevant to a theory of the listener's intuitions.

[3] Extending the Theory of Phenomenal Accents

[3.1] In this section I will explore some of the implications of Lerdahl and Jackendoff's fleetingly proposed connection between performing nuance and perceptual stability in the Mozart excerpt. My approach involves extending one category of accents to the extreme foreground, the level at which performing nuances operate, and it is informed by (mostly) post-*GTTM* investigations of SYVARs in expert performance. I contend that, in the absence of complete metrical stability (i.e., regularity), a class of contextual featuresnamely, features introduced in performancebecomes increasingly relevant to the listener's sense of metrical structure (MS). Here I am simply transposing an argument from Lerdahl's "Atonal Prolongational Structure" from the domain of tonality to that of meter; Lerdahl claims that in music where tonal stability is deliberately compromised, a sense of quasi-tonal prolongation might still be inferred on the basis of contextually salient events.[58]

[3.2] The discussion of MS in *GTTM* begins with an innovative system for classifying the different types of accent cues whose combined effects enable listeners to infer metrical patterns. Lerdahl and Jackendoff's explanation of the relationships between phenomenal and metrical accents

is of crucial importance to the present study, so I shall quote it at some length:

In our judgment it is essential to distinguish three kinds of accents: phenomenal, structural, and metrical. By *phenomenal accent* we mean any event at the musical surface that gives emphasis or stress to a moment in the musical flow. Included in this category are attack points of pitch events, local stresses such as sforzandi, sudden changes in dynamics or timbre, long notes, leaps to relatively high or low notes, harmonic changes, and so forth. By *structural accent* we mean an accent caused by the melodic/harmonic points of gravity in a phrase or sectionespecially by the cadence, the goal of tonal motion. By *metrical accent* we mean any beat that is relatively strong in its metrical context. . . .

Phenomenal accent functions as a perceptual input to metrical accentthat is, the moments of musical stress in the raw signal serve as "cues" from which the listener attempts to extrapolate a regular pattern of metrical accents. If there is little regularity to these cues, or if they conflict, the sense of metrical accent becomes attenuated or ambiguous. If on the other hand the cues are regular and mutually supporting, the sense of metrical accent becomes definite and multileveled. Once a clear metrical pattern has been established, the listener renounces it only in the face of strongly contradicting evidence. . . .

In sum, the listener's cognitive task is to match the given pattern of phenomenal accentuation as closely as possible to a permissible pattern of metrical accentuation; where the two patterns diverge, the result is syncopation, ambiguity, or some other kind of rhythmic complexity. Metrical accent, then, is a mental construct, inferred from but not identical to the patterns of accentuation in the musical surface. [59]

Later, they explain what they mean by "local stresses such as sforzandi" and further explain the relationship between stresses and metrical accents:

By local stress we mean extra intensity on the attack of a pitch-event. We include as signs of stress not only those marked by the signs > and ˆ, but also those indicated by *sf*, *rf*, *fp*, and *subito f*. In a regular sequence of attacked notes, those with stress will preferably be heard as strong beats. [60]

[3.3] The effects of phenomenal accents on metrical accents are formalized as MPRs 4-5, which address relative stress (i.e., loudness) and length, respectively. [61] All the loudness- and length-related accents shown in their exemplars for these rules occur at the rather blatant level that

notation can capture, but I will argue that such accents also occur on a much subtler scale in skilled performance. I will call the latter expressive details *dynamic* and *agogic micro-accents* (DMs and AMs), or collectively, *phenomenal micro-accents* (PMs). In general, their effect on listening should be expected to be rather weak owing to their small scale, but as Lerdahl and Jackendoff demonstrate in the case of the Mozart excerpt, PMs can become extremely important in cases where notational clues offer insufficient support for a single preferred reading. In a sense, AMs seem to be incommensurable with MWFR 4, which states that beats must be evenly spaced. As I mentioned, however, Lerdahl and Jackendoff do acknowledge that even spacing is easily inferred in the case of performances made uneven by expressive timing.[62] I would agree that listeners do seem to "correct" the unevenly spaced beats in expressive performance, but, unlike Lerdahl and Jackendoff, I would draw attention to the fact that important information is communicated in the discrepancy between the actual sounding event and the evenly spaced abstraction.

[3.4] The relationship between metrical accents and what I have called PMs was further clarified in subsequent writings by psychologists John A. Sloboda and Eric Clarke, among others.[63] Whereas Lerdahl and Jackendoff theorized the effects of phenomenal accents on perceived metrical accents, both Sloboda and Clarke studied the effects of perceived meter on PMs. They carried out several experiments to determine the effects of rhythm and meter notation on performance. In one of Sloboda's studies, several melodies were presented to skilled pianists, who performed the melodies on a grand piano monitored by a computer interface. Among the experimental melodies were several that differed only in the placement of the barline; all other parameters, including pitch materials, rhythmic patterns, and expression marks, were unchanged (see Example 7 [DjVu] [GIF]). A similar experiment included melodies that differed only in the time signature. These experiments showed that skilled pianists usually engage a relatively small repertory of SYVARs to project meter, and that their use of these cues is proportional to their degree of experience.[64]

[3.5] Additional empirical studies conducted by both Sloboda and Clarke confirm that experienced listeners can identify the meter intended by the performer in the case of the ambiguous melodies used in their performance experiments, so PMs would seem to be a key to the understanding of each performer's conception of meter in any metrically ambiguous passage.[65] Although the purpose of these experiments was to determine principles for the effects of notation on performance, not to theorize the impact of performance cues on the listening experience, these results are nevertheless relevant to our listener-oriented theory of meter and hypermeter. All that is needed is a reversal of orientation. By

cross-referencing the SYVARs in a given performance with the conventions for projecting meter, we should be able to infer the performer's metrical interpretation of a passage. Among the SYVARs that Clarke and Sloboda identify are dynamic stress and lengthening, which I discussed above in the context of Lerdahl and Jackendoff's MPRs 4–5, as well as the lengthening of the upbeat. [66] I will distinguish downbeat lengthening from upbeat lengthening by referring to the former as "elongation-style" and the latter as "hesitation-style" AMs. Hesitation-style AMs appear to be unrelated to any of the existing MPRs and even seem to contradict MPR 5(a), which states: "Prefer a metrical structure in which a relatively strong beat occurs at the inception of . . . a relatively long pitch-event." [67] Thus, I consider it the weakest of the three classes of PMs relevant to ambiguous cases of MS. Nevertheless, its systematic use is well documented in Sloboda's and Clarke's experiments, and I believe it is widely understood by listeners, so it ought to be reflected in an MPR. Therefore, I will propose the following addition to the Rule Index in *GTTM*: MPR 5.5 (Hesitation): "Weakly prefer a metrical structure in which a relatively strong beat occurs immediately after a relatively long pitch-event." (This rule is included in my Appendix 1, "Revised Rule Index for Metrical Structure.") Assuming that these PM cues (i.e., DMs and the two classes of AMs) might also operate at levels somewhat deeper than surface meter, we now have the theoretical principles needed to assess whether experienced listeners can be expected to infer hypermetric interpretations other than the two predicted by *GTTM* in the case of the opening of Mozart's Symphony No. 40 in G Minor.

[4] Performance Analyses and Transformations

[4.1] A summary of the techniques that I use for performance analysis is included as Appendix 2. Essentially, I convert the desired excerpts into sound files and analyze the timing and loudness with audio editing software. I should emphasize that I have made no attempt at an inductive statistical analysis of these performances, that is, to use raw quantitative data as "input" and propose qualitative judgments as "output." Instead, I prefer to use empirical data selectively (though hopefully not *too* selectively) in order to articulate qualitative PM judgments and comparisons arrived at through careful listening. I begin the performance-analytic procedure with close listening in order to avoid attributing importance to distinctions that are too fine for the ear to detect under normal listening conditions. Quantitative performance analysis is extremely useful in supporting and refining many observations, and it also facilitates detailed inter-performance comparisons. Like other forms of music analysis, quantitative performance analysis seems

helpful in sharpening one's sensitivity to fine nuances—in this case to gradations of intensity and duration.

[4.2] In *GTTM*, beats are considered durationless points in time inferred from the musical stimulus. The acoustical correlate to the beat is the *onset*, or attack point, of a tone that is understood to articulate the beat in question. When we speak loosely of the duration of a beat, we are really talking about the interval *between* beats, and the corresponding acoustical measure is the inter-onset interval (IOI). Similarly, when we speak of the loudness of a beat, we are describing the loudness of some sound within the IOI whose onset corresponds to the beat in question. The acoustical correlate of beat loudness is the peak amplitude within an IOI, also known as the peak sound level (PSL). When a beat is subdivided, the peak amplitude of the first subdivision is used in estimating DMs.

[4.3] The complete data that I collected from the four recordings that I will discuss are given in [Appendix 3](). Caution is often needed in interpreting this numerical information, since the values do not always reflect the listening experience accurately. Sometimes distinctions are so subtle that they are imperceptible under normal listening conditions, so it is important to keep the "just-noticeable differences" in mind. Ballpark figures for these are 5–10% for inter-onset intervals (IOIs) and 0.5–2.0 dB for intensities.[68] Also, both onset perception and intensity perception vary considerably in relation to pitch and timbre.[69] For example, listeners with normal hearing will hear tone x (100 Hz, 50 dB$_{SPL}$) and tone y (1000 Hz, 20 dB$_{SPL}$) as being equal in loudness despite their vast differences in amplitude. Researchers have not seemed to come up with a way to adjust for these perceptual considerations, possibly because of complexities that arise in dealing with multi-voice textures, pedaling, and the interactions of overtones. Also, perhaps most importantly, this type of performance analysis does not allow us to examine the intensities of individual voices, so it is sometimes tempting to misread an intensity analysis as a representation of melodic dynamics rather than the net dynamics of all voices combined. My interpretation of DMs will be based largely on the rankings of downbeat intensities relative to one another, not on absolute values, and, in general, I will draw attention only to those performance analysis statistics that seem most clearly to reflect and enhance the actual listening experience.

[4.4] Example 8 [[DjVu]()] [[GIF]()] includes six hypermetric interpretations of the opening of Mozart's Symphony No. 40. The first two are those that Lerdahl and Jackendoff identify, and the remaining four are drawn from recordings conducted by Benjamin Britten, Neville Marriner, Bruno Walter, and Leonard Bernstein.[70] It is important to note that, although all six representations use *GTTM*-style metrical notation, none is derived from

the strict application of the rules for MS. (Recall that this is the excerpt that Lerdahl and Jackendoff use to illustrate problems in rationalizing hypermetric irregularity and to justify the restrictive MWFR 4, which insists on even spacing of beats and hyperbeats at the tactus and immediately larger metrical levels.)

[4.5] As I mentioned above (see paragraph 2.7), I will account for the irregularities in these metrical patterns by developing transformational rules that apply to MS abstractions. I will consider the eccentric hypermeasure in each version to be a transformation of an underlying regular hypermeasure. Superficially, it might seem that these transformations fracture the metrical structure at this level, resulting in a series of fragments that I will call *metrical structure episodes* (abbreviated MSEs). At level H_1, I submit that the unity underlying each series of MSEs can be understood with little effort on the part of the listener. At deeper levels, however, an underlying unity cannot always be demonstrated. Nevertheless, in the interest of theorizing a rather subtle aspect of the listening experience, I think it is worthwhile to attempt to show the integrity of each independent MSE, rather than vaguely stating (as Lerdahl and Jackendoff do) that irregularities cause MS to become attenuated at deeper levels. Indeed, this approach allows us to trace in detail the gradual breakdown of MS from level to level.

[4.6] Let us begin with the two readings proposed by Lerdahl and Jackendoff (Example 8a–b). Each of these examples includes one instance of triple meter in the context of a prevailing duple meter at the first level of hypermeter. The eccentric hypermeasure might be understood to result from the process of metrical expansion, the addition of a second weak beat between two strong beats. In order to reveal the regular underlying structure, a rule for the opposite process is required, a process which I will call "metrical contraction." This rule should simply state that, in order to regularize a pattern such as this, one of the weak beats in the three-beat measure must be deleted. This rule can be stated formally as follows:

Metrical Contraction[71]

Given
(i) a well-formed metrical structure episode M that ends with beats B_1 and B_2, in which B_1 and B_2 are adjacent beats at level L_i and B_1 is also a beat at level L_{i+1}, and
(ii) a well-formed metrical structure episode N in which B_3, B_4, and B_5 are adjacent beats at level L_i and only B_3 is also a beat at level L_{i+1}, and
(iii) a well-formed metrical structure episode P that begins with beats

B_6, B_7, and B_8, in which B_6, B_7, and B_8 are adjacent beats at level L_i and both B_6 and B_8 are also beats at level L_{i+1},
and given that M, N, and P are adjacent metrical structure episodes,

then a well-formed metrical structure episode M' can be formed by deleting B_5, such that B_1, B_2, B_3, B_4, B_6, B_7, and B_8 are adjacent beats at level L_i and B_1, B_3, B_6, and B_8 are also beats at level L_{i+1}.

[4.7] In a performance that projects Interpretation A (Example 8a), the listener might at first perceive three MSEs in mm. 1–23: a series of five duple hypermeasures (mm. 1–10), followed by one triple hypermeasure (mm. 11–13), followed by five more duple hypermeasures (mm. 14–23). If Interpretation B (Example 8b) is projected, the listener would instead perceive three duple hypermeasures (mm. 1–6), followed by one triple hypermeasure (mm. 7–9), followed by seven duple hypermeasures (mm. 10–23). In both cases, the listener might then intuitively reconceive the entire passage as a coherent sequence of eleven duple hypermeasures, as suggested by the Metrical Contraction rule (Example 9a–b).[72] That is not to say that the differences between the two performances will be ignored, but rather that the two performances will be understood as departing from the same underlying metrical structure in different ways.

[4.8] Two of the remaining four interpretations (Example 8c and f [DjVu] [GIF]) can be understood to show metrical truncations rather than expansions. That is, in the context of a prevailing duple meter at level H_1, a new hypermeasure begins before the preceding one has been completed, such that two strong beats are found side-by-side. In order to reverse this process, we require a transformational rule for "metrical completion," which will stabilize the MS at this level by inserting a weak beat between these strong beats. This rule can be stated formally as follows:

Metrical Completion

Given
(i) a well-formed metrical structure episode M that ends with beats B_1, B_2, and B_3, in which B_1, B_2, and B_3 are adjacent beats at level L_i and both B_1 and B_3 are also beats at level L_{i+1}, and
(ii) a well-formed metrical structure episode N that begins with beats B_4, B_5, and B_6, in which B_4, B_5, and B_6 are adjacent beats at level L_i and both B_4 and B_6 are also beats at level L_{i+1},
and given that M and N are adjacent metrical structure episodes,

then a well-formed metrical structure episode M' can be formed by inserting beat B_x between beats B_3 and B_4, such that B_1, B_2, B_3, B_x, B_4, B_5,

and B_6 are adjacent beats at level L_i and B_1, B_3, B_4, and B_6 are also beats at level L_{i+1}.

[4.9] In all four recordings, strong and weak hyperbeats alternate regularly at level H_1 in mm. 1–10, and m. 20 has a strong accent, so I will focus on what happens in mm. 11–20. The shift occurs earliest in Britten's version through a succession of strong accents on the downbeats of both m. 13 and m. 14 (see Example 8c). The accent at 13.1 relative to 12.1 is achieved through a hesitation-style AM (in this case, an extension of IOI 12.2 by 14.8%) and a DM (an increase in net amplitude by 0.9 dB and three ranking points at the downbeat-to-downbeat level).[73] But 14.1 is stronger still (by nearly 4 dB and two ranking points), and 15.1 sounds softer than 14.1 despite the increase in orchestration. (The decrease in net amplitude is insignificant at 0.1 dB, but, in light of the expanded texture, the absence of an increase in amplitude supports a strong-weak hypermetric pattern in mm. 14–15.) A diminuendo added by Britten through m. 15, an abrupt change in orchestration at 16.2, and an absence of clear AMs in these measures create some confusion, but the emphasis on even-numbered downbeats is confirmed in mm. 17–20. IOI 18.1 has an elongation-style AM (9.2% longer than 17.1), and 19.1 is markedly softer than 18.1 as well (by a margin of 1.8 dB). To further reinforce the even-measure accents, 20.1 has an elongation-style AM (7.5%) and is the loudest beat in the entire excerpt, which seems especially dramatic in light of the absence of a crescendo in m. 19. The resulting pattern has adjacent strong beats at 13.1 and 14.1, indicative of a metrical truncation, so the passage can be thought of as containing two MSEs (mm. 1–13 and 14–20), and a regular underlying pattern can be generated by applying the Metrical Completion rule ([Example 10](), Audio: [[Modem]]([Broadband]]).

[4.10] In contrast to Britten's interpretation, Marriner preserves the odd-measure emphasis until the last possible moment, mm. 19–20 (see Example 8d [DjVu] [GIF]). In mm. 10–20, this is projected extremely clearly through elongation-style agogic accents at 11.1, 13.1, 15.1, and 17.1 (by 20.0%, 10.9%, 9.4%, and 41.7%, respectively, at the downbeat level) as well as dynamic accents on these same beats (all of which have higher rankings than the even-measure downbeats that surround them). Indeed, 17.1 rather than 20.1 is the loudest downbeat in the excerpt. The arrival of even-measure accentuation at 20.1 seems to be associated mostly with the surface rhythm's agogic accent, which is further enhanced with an elongation-style AM (6.4%). Note also that 19.1 is the first weak odd-measure downbeat in the excerpt; it is somewhat quieter than 18.1 (−1.7 dB) and about equal in length (IOI 19.1 is only 2 ms or 0.4% longer than 18.1, an imperceptible difference). The adjacent weak beats at 18.1

and 19.1 constitute a metrical expansion, and the pattern can be regularized through the application of the Metrical Contraction rule ([Example 11](), Audio: [[Modem]()] [[Broadband]()]).

[4.11] Walter's hypermetric interpretation is more complex than the preceding two, and this is largely because of clever ambiguities in his deployment of PMs (see Example 8e [[DjVu]()] [[GIF]()]). The odd-measure strong beats established in mm. 1–10 continue at least through m. 13. 11.1 is both longer (10.2%) and louder than 10.1, and the relation of 13.1 to 12.1 is similar (8.7% longer, 1.2 dB louder overall).[74] It is difficult to offer a hypermetric interpretation of mm. 14–16, because the PMs play against the notated surface meter. Performers sometimes refer to this effect as a "negative accent" or "deflection," and it occurs when a downbeat is considerably quieter than the listener would expect on the basis of the upbeat. (See the final column of [Appendix 3c](). These are the only downbeats in the entire excerpt that have negative changes in intensity at the beat-to-beat level.) Without a clear projection of surface meter, it is difficult to assess the location and strength of hypermetric accents. In the absence of evidence to the contrary, I would be inclined to hear a continuation of the odd-measure accents through to m. 16. It could be argued that the slight, though noticeable, acceleration into 15.1 supports this reading by giving that beat a special emphasis. (IOI 14.2 is shortened by 11.2%.) Next, two features conspire to encourage us hear a continuation of the odd-measure accents at 17.1: the sudden increase in dynamics and orchestration at 16.2 and the literal repetition of 16.2–17.1 in 17.2–18.1. All else being equal, 17.1–18.1 would most likely be heard as a strong-weak echo effect ([Example 12a](), Audio: [[Modem]()] [[Broadband]()]), a reading consistent with MPR 2.[75] Walter neither underlines nor contradicts this reading, however. The first three downbeats after 16.2 sound equal in loudness (0.4 dB difference), and despite some elongations that enhance the syncopation effect on the second beat of each measure, mm. 17–19 sound steady. (All three downbeats are within 5% of the average tempo of the entire excerpt.) Thus, it could be argued that Walter leaves the hypermetric interpretation undefined in mm. 17–19. There is, however, a salient accent at 20.1 (loudest downbeat IOI in the excerpt, 1.4 dB louder than 19.1), and in retrospect this might lead us to hear the entire cadential extension (mm. 16–20) in the context of an even-measure hypermeter (Example 12b). At that point we would realize that our hunch that 17.1–18.1 was a strong-weak echo was incorrect. In that sense, the metric and hypermetric ambiguities of mm. 14–19 (especially mm. 17–19) in the Walter recording have a rather humorous effect, and one that adds richness to the listening experience. The reading shown in Example 12b (and in Example 8e) is the final version that emerges once the accent at 20.1 is heard. Walter's interpretation includes

a metrical expansion, this time in mm. 15–17, so the pattern can be regularized through the Metrical Contraction rule (Example 12c).

[4.12] Bernstein's hypermetric interpretation (Example 8f [DjVu] [GIF]) is nearly identical to Marriner's (Example 8d) at level H_1. The most important distinction is that in Bernstein's recording, there are strongly pronounced PMs at both 19.1 (2.6 dB louder than IOI 18.1) and 20.1 (20.8% elongation, 1.2 dB increase compared to 19.1). Thus, 19.1 is a strong hyperbeat, like all the preceding odd-measure downbeats, rather than a weak hyperbeat, like 19.1 in Marriner's recording. In the case of Bernstein's recording, the pattern can be regularized through the Metrical Completion rule (Example 13 Audio: [Modem] [Broadband]). At other points in the excerpt, Marriner and Bernstein project the same interpretation by quite different means. For instance, Bernstein's accent at 11.1 seems to be communicated through timbre rather than loudness and elongation.[76]

[4.13] So far I have been describing only the first level of hypermeter, that is, MS at the level of the double-measure. Lerdahl and Jackendoff suggest that a deeper level of hypermeter should include strong accents at 16.1 and 22.1, because these are points of harmonic arrival and might therefore be understood as structural accents. Stated in the terminology I have developed, this means that mm. 16–21 form an MSE containing a three-beat hypermeasure plus a downbeat at level H_2 (Example 14 [DjVu] [GIF]). Although Lerdahl and Jackendoff situate this six-bar episode at the "4-bar level," implying that it is a transformation of two underlying two-bar MSEs, they do not clarify this transformation because they rightly consider it futile to attempt such a thing in music-notational terms and on the basis of the score alone.[77] Overall, Lerdahl and Jackendoff consider level H_2, aside from this singular MSE, to be too ambiguous to explore, and they even go so far as to say, "The 4-bar level simply does not have much meaning for this passage."[78] I would argue that, despite the ambiguity of the score, metrical organization across spans of four or even eight measures can be conveyed, at least episodically, in the performance of this excerpt. Owing to the instability caused by the erosion of regularity at these levels, the MS interpretation becomes increasingly dependent on salient PMs. (Here I am again invoking Lerdahl's hypothesis that contextual salience becomes crucial to the construction of hierarchical structures when stability conditions are compromised. See paragraph 3.1 above.) Although the structures I am about to describe lie deep within the grey area between stability and instability, or between comprehensibility and incomprehensibility, I think they deserve some consideration.

[4.14] In Britten's recording, an MSE consisting of an upbeat plus two duple hypermeasures can be shown at the level H_2 in mm. 1–10 (Example 15a [DjVu] [GIF] Audio [Modem] [Broadband]). This can be confirmed by comparing the beats found at the level H_1 within this passage (i.e., beats 1.1, 3.1, 5.1, 7.1, and 9.1). Of these, 1.1 is clearly the weakest. (Indeed, it is ranked last among the first twenty downbeats.) A strong-weak pattern is then established through qualitative differences between 3.1 and 5.1. Whereas the former is prepared with a hesitation-style AM (IOI 2.2 is 7.1% longer than 2.1), the latter is anticipated slightly (IOI 4.2 is 8.8% shorter than 4.1), and this conveys the impression that 3.1 is accented with conviction, while 5.1 occurs almost by accident. Similarly, the DM on 5.1 sounds like a "jolt" in all voices, while 3.1 has a distinct melodic accent.[79] Still clearer is the distinction between 7.1 and 9.1. The former is both louder and longer than the latter (by a margin of 5.2 dB and 9.2%, respectively), and at the measure-to-measure level, 7.1 has a stronger elongation-style AM and a stronger DM than 9.1 (20.8% vs. 5.1%, and 5.3 dB vs. 1.5 dB, respectively). A further distinction can be shown between the two strong beats at the level H_2, and this produces a third level of hypermeter in this MSE. Beat 7.1 has a greater DM and AM than 3.1, and this creates an upbeat-downbeat figure (Example 15b). Beats 3.1 and 7.1 are preceded by similar degrees of hesitation (7.1% and 6.3%, respectively), but 7.1 has an additional elongation-style AM (13.6%) as well as a considerably stronger DM than 3.1 (5.3 dB at the downbeat-to-downbeat level, vs. 2.5 dB). Note, however, that this interpretation is rather fragile. It includes only a partial hypermeasure, and the downbeat's position in the seventh measure of a ten-measure MSE seems unusually late. The latter objection is formalized in GTTM as MPR 2 (Strong Beat Early): "Weakly prefer a metrical structure in which the strongest beat in a group appears relatively early in the group."[80]

[4.15] After the conclusion of the first MSE in Britten's recording, hypermeter seems to be attenuated at level H_2 until m. 14. This is because each of the next three strong beats at level H_1 (i.e., 11.1, 13.1, and 14.1) has a stronger DM than the last (+0.5 dB, +0.6 dB, and +3.8 dB, respectively). A second MSE (Example 16a [DjVu] [GIF] Audio [Modem] [Broadband]) is weakly suggested, however, by PMs in mm. 14–20. The distinctions are again largely qualitative. The DM at 20.1 (+2.3 dB compared to IOI 19.1), in combination with the changes in texture, harmony, and dynamics that immediately follow it, seems more prominent than the elongation-style AM at 18.1 (12.4%). Also, as discussed above, the status of 16.1 as a beat at this level is achieved mainly by inference, in light of the complications introduced by the change in texture at 16.2. It would therefore seem reasonable to regard it as a weaker beat than 14.1, which is very strongly stressed (+3.8 dB compared to IOI 13.1) and also

considerably stronger than 16.1 (by a margin of 6.1 dB). I will not attempt to continue this MSE beyond m. 21, because the next potential beat, 22.1, also has an accent (a structural accent, in this case), and two strong beats cannot be adjacent in an MSE. Of the four beats in this MSE, the first and last are most strongly accented, so mm. 14–19 can be considered to form a three-beat hypermeasure at level H_2 (Example 16a). Although it is tempting to consider this an expansion of the duple pattern found in the first MSE, this interpretation lacks a sufficient metrical context according to the criteria I have established, since the second MSE is not *directly* preceded by a duple MSE.[81] A comparison of the two strong beats at H_2 in this MSE (i.e., 14.1 and 20.1) reveals a sufficient distinction to propose a third level of hypermeter here. As in the first MSE, the second of these two hyperbeats is stronger than the first on the basis of both AM and DM cues, and this yields an upbeat-downbeat figure at H_3. (IOI 20.1 has 8.7% elongation and is ranked first among the even downbeats in mm. 1–20, while 14.1 has 3.1% elongation and is ranked sixth.)

[4.16] The remaining three recordings evince more stable MSEs than Britten's at the second and third hypermetric levels, and two of them (Marriner's and Bernstein's) include some tenable hypermetric transformations at the level H_2. In Marriner's recording, the first MSE continues through m. 16 (Example 17a [DjVu] [GIF] Audio [Modem] [Broadband]). Up to m. 10, it projects the same metrical pattern as Britten's recording (see Example 14 [DjVu] [GIF]), but in Marriner's version the alternation of strong and weak hyperbeats continues through 15.1. In terms of overall intensity, IOI 11.1 is louder than 9.1 (+7.6 dB) and likewise IOI 15.1 is louder than 13.1 (+8.1 dB). The long crescendo in mm. 12–15 adds further emphasis to 15.1, but it also obscures the relationship of 11.1 and 13.1 somewhat; nonetheless, 11.1 has a considerably longer elongation-style AM than 13.1 (35%, vs. 13%), so a strong-weak pattern is projected here. In Marriner's recording, the first MSE is twice as long as in Britten's recording, so we might expect that the third hypermetric level is somewhat more stable. Indeed, the hyperbeats at this level (i.e., 3.1, 7.1, 11.1, and 15.1) do evince a weak-strong-weak-strong pattern on the basis of their DMs (ranked sixth, fourth, seventh, and third, respectively, among odd-measure downbeats in mm. 1–20). The resulting pattern can be interpreted as an upbeat, a full hypermeasure, and a downbeat (Example 17b [DjVu] [GIF]). The first MSE ends at m. 16, because 17.1 is a strong beat (indeed, it has the loudest IOI in the excerpt) rather than the weak beat that we had come to expect. This strong beat belongs to another MSE at this level, which begins with a strong-weak-strong pattern among the next three hyperbeats, 17.1, 20.1, and 22.1 (Example 17c). As I mentioned above, the loudest downbeat in the passage is 17.1 in Marriner's recording, and 22.1 is a strong beat at this

level because of the structural accent articulated by the return of tonic harmony. Because these metrical accents are caused by two different categories of accent cues, phenomenal and structural, it would be arbitrary to identify one as being stronger than the other, and furthermore I don't think either alternative would have much intuitive appeal. For both these reasons, I think it would be artificial to propose a third level of hypermeter for this MSE. It does, however, seem possible to combine the two MSEs at the second level into a single MSE, since they are adjacent and each contains at least one full duple measure. Because there is a sequence of two strong beats at their border (i.e., 15.1 and 17.1), we must use the Metrical Completion rule in order to regularize this pattern (Example 17d).

[4.17] Walter's interpretation is arguably the most stable of the six under consideration. On first hearing, we might consider the first MSE at H_2 to end in m. 10, as in Britten's recording. A pattern of alternating weak and strong hyperbeats seems to be established through DMs in 1.1, 3.1, 5.1, 7.1, and 9.1 (ranked tenth, fifth, ninth, third, and eighth among the odd-measure downbeats in mm. 1–20), and this pattern seems to be disrupted by the presence of another weak beat at 11.1 (5.4 dB quieter than IOI 13.1). If we consider 1.1 a strong beat, however, on the basis of its early position (see MPR 2), then a triple-meter pattern is established by m. 7, and this pattern continues uninterrupted through the entire excerpt (Example 18a [DjVu] [GIF] Audio [Modem] [Broadband]). According to the latter interpretation, the first MSE of the piece ends at m. 21, because the structural accent at 22.1 would otherwise yield two adjacent strong beats (i.e., 20.1 and 22.1). The four strong beats in this MSE (1.1, 7.1, 13.1, and 20.1) are established as follows: 1.1 is strong because of its relatively early position (MPR 2); 7.1, because it is louder than the surrounding hyperbeats (5.1, 7.1, and 9.1 are ranked ninth, third, and eighth); 13.1, because it is louder than the preceding hyperbeat (+5.4 dB) and has a greater elongation-style AM than the ensuing one (15.9%, vs. 6.0%); and 20.1, because it is the loudest beat in the entire excerpt. On the basis of both DMs and AMs, beats 7.1 and 20.1 are the strongest of these four beats (1.1, 7.1, 13.1, and 20.1 rank twentieth, fifth, eighth, and first in loudness among the twenty downbeats in the excerpt, and IOI 7.1 has the most pronounced AM in the excerpt, an elongation of 30.4% over IOI 6.1.) This creates a weak-strong-weak-strong pattern at H_3 (Example 18b). We might tentatively posit a fourth level here as well. It is difficult to choose between 7.1 and 20.1 on the basis of PMs, because 7.1 has the longest IOI in the excerpt and 20.1 the loudest. Nevertheless, the contrasts in texture, dynamics, and harmony that immediately follow 20.1 make it seem the more salient of the two beats, so an upbeat-downbeat figure seems to be the more defensible choice (Example 18c). As we found

in both MSEs of Britten's recording at level H_3, the absence of a complete hypermeasure, combined with the extremely late position of the strong beatthis time at the downbeat of the twentieth measure out of twenty-one!makes the H_4 interpretation here extremely fragile.

[4.18] As I mentioned previously, Bernstein's interpretation is nearly identical to Marriner's at level H_1; they differ only in the nature of the transformation immediately preceding 20.1. As we shall see, however, the two interpretations diverge considerably at subsequent levels of MS. In Bernstein's version, the first MSE continues only to m. 10, because (as in Britten's recording) weak beats at level H_2 are found at both 9.1 and 11.1. The five beats in this MSE (i.e., the odd-measure downbeats in mm. 1-10) are differentiated in a way that should by now seem familiar, that is, through an alternation of relatively weak and relatively strong DMs (Example 19a [DjVu] [GIF] Audio [Modem] [Broadband]). (Beats 1.1, 3.1, 5.1, 7.1, and 9.1 are ranked tenth, fifth, seventh, fourth, and ninth among the odd-measure downbeats in mm. 1-20.) Of the two strong beats in this MSE, 7.1 is the stronger, for it has a more pronounced AM and DM than 3.1. (IOI 7.1 has 11.2% elongation at the beat level, vs. 5.9% in IOI 3.1, and the PSL in IOI 7.1 is 2.8 dB louder than in IOI 3.1). Thus, the first MSE has an upbeat-downbeat figure at H_3, much like the one found in Britten's recording (Example 19b). The weak beat at 11.1 initiates a second MSE (mm. 11-21), which also has an alternation of weak and strong hyperbeats (Example 19c). The differentiation of 11.1 and 13.1 is clear on the basis of their DMs (ranked fourteenth and eighth among the downbeats in mm. 1-20), as is the differentiation of 15.1, 17.1, 19.1, and 20.1 (ranked fifth, first, third, and second). Also note that, although IOI 13.1 is quieter than 15.1, the latter is less strongly delineated at the beat-to-beat level; whereas IOI 13.1 is louder than 12.2 (+1.0 dB), 15.1 is somewhat quieter than 14.2 (-0.8 dB). Among beats 13.1, 17.1, and 20.1, the three strong beats in the second MSE, the one with the loudest IOI is 17.1, so a weak-strong-weak pattern might be inferred at level H3 (Example 19d).

[4.19] Both MSEs in Bernstein's recording have duple meter at level H_2, and two weak beats are adjacent at the border between the two. This interpretation can therefore be transformed to a regular, well-formed MSE spanning the entire excerpt through the Metrical Contraction rule (Example 20a). This transformation also yields a well-formed MSE at level H_3, consisting of an upbeat and two full duple measures (Example 20b). Although these two hypermeasures are conceptually equal in length, the spans of the musical surface to which they correspond (i.e., mm. 7-16 and 17-21) are radically different in lengththey cover ten and five measures, respectively. (Incidentally, this progressive shortening is also evident at the level of the hyperbeats in these measures, which span six, four,

three, and two measures.) This asymmetry might influence our choice between 7.1 and 17.1 at level H_4. Beat 17.1 is clearly the louder of the two (by a margin of 13.2 dB), and it also has a salient hesitation-style AM (10.5% at the beat-to-beat level), while 7.1 has an elongation-style AM (11.2%). Although 17.1 would appear to be the stronger beat on the basis of these PM cues, this does not seem to be an intuitively justifiable reading. It seems that the extremely wide spacing of these hyperbeats diminishes the force of PM comparisons. Instead, 7.1 seems like the stronger beat (see Example 20c), and this reading is supported by the asymmetry between the portions of the musical surface corresponding to the two hyperbeats (see Appendix 3, MPR 5a) and also by the tendency to hear the earlier of two more-or-less equally accented beats as the strong beat (MPR 2).

[4.20] As we have seen, each of the six interpretations under consideration employs a different set of hypermetric transformations and thus conveys a different version of the work. Some might find the diversity among these six interpretations unsettling, and might consider the absence of criteria for determining which is the "correct" or "intended" version to be a shortcoming of the procedure I have developed. To this objection, I would respond that, like the Gestalt theorists and like Lerdahl and Jackendoff, I am interested in understanding structural intuitions without recourse to idealism. In focusing on general principles of perception and on conventions for the projection of meter, I have attempted to sidestep the problematic metaphysical assumption that the work itself is timeless and fully determined prior to any performance of it, an assumption that is often implicit in critical discourse on the relative merits of different interpretations. Instead of imagining an ideal performance, I prefer to use hypothetical, metrically regular abstractions inferred from the interaction of score and performance as the framework for comparing interpretations. In order to use these comparisons to support an argument about the merit of a recording, we would need to postulate specific criteria for critical evaluation. For instance, if we decide that symmetry is important, then we might argue that Britten's performance is outstanding because of the parallelism between the two MSEs in his version of the Mozart excerpt at levels H_2 and H_3 (see Examples 15 [DjVu] [GIF] and 16 [DjVu] [GIF]). If we value musical humor, then we might instead prefer Walter's interpretation, because of the thwarted expectancy at level H_1 in his recording (see Example 12). While this approach to criticism might be interesting, it is important to realize that the comparisons enabled by my adaptations to the Lerdahl-Jackendoff theory are, in themselves, non-judgmental.

[5] Rethinking the Role of Performance in the Lerdahl-Jackendoff Theory

[5.1] If Lerdahl and Jackendoff are correct in characterizing the Mozart example discussed in Part 4 as "a not untypically complex passage" with regard to Metrical Structure, then it would appear that their "balance-tipping" analogy, which I quoted previously (see paragraph 2.4), is in need of some refinement. The expression "tip the balance" implies that performance-specific features are considered relatively late in the analytic process, if at all, after an impasse is reached in the interpretation of a score. That is, in conducting an analysis, we first study the score, and in doing so we discover something that resembles the faces/vase illusion, so we then listen to a recording, and finally decide on a preferred reading. Because the Lerdahl-Jackendoff theory is intended first and foremost to depict the *listener's* intuitions, I believe this sequence of events is inappropriate. By beginning with the details of performances in an ambiguous passage, I have shown that more than two different interpretations can be conveyed, and that these interpretations are not necessarily the ones that are the most obvious on the basis of the score. The potential multivalence of a musical structure is sometimes more extravagant than a score-reader can anticipate; some passages are more like mosaics than faces/vase illusions in the range of meanings they can evoke (see Example 1b, j [DjVu] [GIF]). I therefore suggest that we begin our revision to the schematic representation of the "balance-tipping" model by extending the number of possible interpretations indefinitely (Example 21 [DjVu] [GIF]).

[5.2] To me, the score-reader's inability to predict a performer's interpretation in no way indicates a lack of musical intuition, but instead reflects elite performers' uncanny ability to avoid sounding predictable, an ability that involves a combination of intuition and conscious thought. *GTTM* goes a long way toward formalizing intuition, but it makes no allowance for the influence a performer's unpredictable, conscious decisions can have on an experienced listener's intuitions-indeed, Lerdahl and Jackendoff regard the performer's construction of an interpretation as a "largely unconscious" process.[83] An influential study by psychologist Caroline Palmer provides a different perspective. In a series of experiments on the performance of works with ambiguous phrase structures, Palmer examines correlations between pianists' score annotations and SYVAR patterns that are conventionally used to communicate interpretations of phrase structure.[84] Each pianist's annotations suggest a different structural interpretation, and that

interpretation corresponds strongly and uniquely to SYVARs in that individual's performance. Palmer's experiment suggests a model of the relationship between ambiguous notation and the sounding musical surface (see Example 22a [DjVu] [GIF]) that is quite different from, but not incompatible with, the model implied by Lerdahl and Jackendoff through their balance-tipping analogy (see Example 6b-c [DjVu] [GIF]). Palmer's model allows maximum room for the performer's freedom of conscious interpretation, and thus it sheds light on exceptions to Lerdahl and Jackendoff's assumption that performing nuance arises largely from unconscious interpretation. Thus, we might further revise our schematic representation of the balance-tipping phenomenon by adding conscious interpretation as an intermediate stage between the intuitive understanding of hierarchical structures and the application of SYVARs (Example 22b [DjVu] [GIF]). A limited form of conscious thought (i.e., preference between two equally legitimate options) was already implicit in Example 4 [DjVu] [GIF], but Palmer's model allows us to expand this considerably to include whatever musical (and even extramusical) considerations seem relevant.

[5.3] Throughout this paper, I have attempted to follow Lerdahl and Jackendoff's lead in theorizing aspects of comprehensibility and remaining silent on aesthetic issues such as, for instance, the thwarting of expectancies or the representation of affect, motion, or aspects of a narrative. Although I would not deny that many of the SYVARs that expert performers use are based on aesthetic considerations such as these, rather than the straightforward projection of an intuitive, stable structural interpretation, I nevertheless see the value in examining aesthetically motivated idiosyncrasies in contradistinction to norms associated with intelligibility. In the realm of linguistics, Noam Chomsky theorized grammatical norms under the rubric of "competence" and referred to the relationship between actual utterances and these norms, including errors and ambiguities, as "performance."[85] In its inclusion of a place for performers' diverse conscious interpretative decisions, many of which (hopefully) will be motivated by aesthetic considerations rather than purely matters of musical syntax, the schematic model I have developed might be useful in the development of a theory of musical "performance," in Chomsky's sense of the term.[86]

[5.4] If *GTTM* were designed as a reading theory rather than a listening theory, or if the Mozart example possessed an exceptionally high degree of metrical complexity, then it might be argued that the scheme I am suggesting places too much emphasis on seemingly superficial performing nuances at the expense of more powerful score-based elements. I would answer this objection by reiterating three crucial facts: (1) the theory

is clearly intended to capture the experienced listener's intuitions, so the aural stimulus rather than the score is the more suitable object of inquiry, (2) Lerdahl and Jackendoff suggest that the Mozart example has an unexceptional degree of metrical complexity, and (3) elsewhere Lerdahl suggests that salience conditions become increasingly important to the listener's construction of hierarchical interpretations in cases where stability is compromised, and in my view this should include violations to MWFR 4, the rule that insists on evenly spaced beats at the tactus and immediately larger metrical levels. I would also point out once again that, in simpler metrical situations that adhere to MWFR 4 at the first few levels of hypermeter, skilled performers and (other) readers should be expected to agree on their interpretations of the score. In the case of an unambiguous structure, therefore, the differences among performances become less relevant, and the score can adequately serve as an substitute for a conventional performance (see Example 6a [DjVu] [GIF]).

[5.5] I will conclude by mentioning a few ways in which the concepts I have discussed might be developed further. Under close examination, recordings of additional hypermetrically ambiguous passages would no doubt prompt refinements to the metrical transformation rules that I proposed in Part 4. The utility of the metrical transformation rules at the level of local meter might also be assessed, using examples from non-Western and twentieth-century repertories in which temporal organization lies in the grey area between the random and the perfectly regular (e.g., changing meters). It is also important to remember that Metrical Structure is one of four interdependent parameters of hierarchical organization theorized in *GTTM*, and it remains to be seen whether the metrical transformations I have described would have effects on the other parameters.[87] It might also be interesting to explore the effects of performing nuance within each of the other parameters of the theory. In particular, the projection of phrase-periodic structure and, more recently, the communication of patterns of musical tension in performance have been studied extensively by music psychologists, and this research would inform an examination of the role of performance in the communication of phrase-periodic structure and patterns of tension and relaxation, parameters that Lerdahl and Jackendoff cover under the rubric of Grouping Structure (GS) and Prolongational Reduction (PR), respectively.[88] The techniques for SYVAR analysis that I outline in Part 4 might also help in the development of analytical applications for the copious insights on performance that are found in more recent texts on meter, such as Jonathan Kramer's *The Time of Music*, William Rothstein's *Phrase Rhythm in Tonal Music*, and Christopher Hasty's *Meter as Rhythm*. Indeed, any structuralist theory that purports to illuminate the listening experience should have room for considering the impact of

performing nuance, and performance analysis seems especially useful in exploring the inner workings of ambiguous structures.

[6] Appendix 1. Revised Rule Index for Metrical Structure⁽⁸⁹⁾

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

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

MWFR 3 [90]
At each metrical level, strong beats are spaced either two or three beats apart.

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

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

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

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

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

MPR 5 (Length)
Prefer a metrical structure in which a relatively strong beat occurs at the inception of either
a. a relatively long pitch-event,
b. a relatively long duration of a dynamic,
c. a relatively long slur,

d. a relatively long pattern of articulation,
e. a relatively long duration of a pitch in the relevant levels of the time-span reduction, or
f. a relatively long duration of a harmony in the relevant levels of the time-span reduction(harmonic rhythm).

MPR 5.5 (Hesitation) (paragraph 3.5)
Weakly prefer a metrical structure in which a relatively strong beat occurs immediately after a relatively long pitch-event.

MPR 6 (Bass)
Prefer a metrically stable bass.

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

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

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

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

Metrical Deletion
Given a well-formed metrical structure M in which
i. B_1, B_2, and B_3 are adjacent beats of M at level L_i, and B_2 is also a beat at level L_{i+1},
ii. T_1 is the time-span from B_1 to B_2 and T_2 is the time-span from B_2 to B_3, and
iii. M is associated with an underlying grouping structure G in such a way that both T_1 and T_2 are related to a surface time-span T' by the grouping transformation performed on G of
(a) left elision or (b) overlap,
then a well-formed metrical structure M' can be formed from M and associated with the surface grouping structure by
(a) deleting B_1 and all beats at all levels between B_1 and B_2 and associating B_2 with the onset of T', or

(b) deleting $B2$ and all beats at all levels between B_2 and B_3 and associating B_1 with the onset of T'.

Metrical Contraction (paragraph 4.6)
Given
(i) a well-formed metrical structure episode M that ends with beats B_1 and B_2, in which B_1 and B_2 are adjacent beats at level L_i and B_1 is also a beat at level L_{i+1}, and
(ii) a well-formed metrical structure episode N in which B_3, B_4, and B_5 are adjacent beats at level L_i and only B_3 is also a beat at level L_{i+1}, and
(iii) a well-formed metrical structure episode P that begins with beats B_6, B_7, and B_8, in which B_6, B_7, and B_8 are adjacent beats at level L_i and both B_6 and B_8 are also beats at level L_{i+1},
and given that M, N, and P are adjacent metrical structure episodes, then a well-formed metrical structure episode M' can be formed by deleting B_5, such that B_1, B_2, B_3, B_4, B_6, B_7, and B_8 are adjacent beats at level L_i and B_1, B_3, B_6, and B_8 are also beats at level L_{i+1}.

Metrical Completion (paragraph 4.8)
Given
(i) a well-formed metrical structure episode M that ends with beats B_1, B_2, and
B_3, in which B_1, B_2, and B_3 are adjacent beats at level L_i and both B_1 and B_3 are also beats at level L_{i+1}, and
(ii) a well-formed metrical structure episode N that begins with beats B_4, B_5, and B_6, in which B_4, B_5, and B_6 are adjacent beats at level L_i and both B_4 and B_6 are also beats at level L_{i+1},
and given that M and N are adjacent metrical structure episodes, then a well-formed metrical structure episode M' can be formed by inserting beat B_x between beats B_3 and B_4, such that B_1, B_2, B_3, B_x, B_4, B_5, and B_6 are adjacent beats at level L_i and B_1, B_3, B_4, and B_6 are also beats at level L_{i+1}.

[7] Appendix 2. Performance Analysis Techniques

[7.1] Before I begin explaining some basic techniques for the analysis of duration and loudness in recordings, I must once again emphasize that the process should begin with close listening unmediated by a computer. This helps us avoid attributing importance to distinctions that are too fine for the ear to detect under normal listening conditions.

[7.2] When a performer is present as an experimental subject, a computer interface with the instrument can allow the researcher to gather data very

efficiently, but data collection is rather more arduous in the case of a recorded performance. At present, there are three basic options for the analysis of recordings: tapping software, spectrographic analysis, and audio editing software (sometimes referred to as waveform analysis). In the first of these approaches, the listener taps along with the perceived beat of the performance, from which the computer calculates rough data on tempo fluctuation.[91] This is the most efficient method for gathering data, and the only realistic method for analyzing complete large-scale works or movements or for making general comparisons of a large number of recordings. However, it is limited to the parameter of tempo, and even in this capacity it is considerably less accurate than the other forms of analysis. Spectrographic analysis, the second approach, is especially useful for gathering information on timbre.[92] This is the preferred method for dealing with orchestral instruments and the human voice (in both live and recorded performances), but it requires cumbersome (and expensive) hardware and a greater knowledge of acoustics and mathematics than many music theorists possess (including the present author). This leaves waveform analysis, in which excerpts from an LP or a CD are converted to digital sound files (e.g., .au or .wav) and analyzed with audio editing software.[93] This method is more accurate than the tapping approach and more user-friendly than the spectrographic, and it can provide reliable information pertaining to loudness and (especially) duration.

[7.3] Once the excerpt has been converted to a sound file, its waveform can be displayed in an audio editing program, and playback can be initiated from any point, with a resolution greater than 1 ms (millisecond).[94] (I prefer Syntrillium CoolEdit, a Shareware program that I have found to be reliable and user-friendly. WaveLab and ProTools are other popular options.) Through a combination of visual and aural observation, one can identify the temporal location of the onset (beginning) of any event, such as a solid chord or an unaccompanied melody note. By using this method, the durations of inter-onset intervals (IOIs) can be calculated at the level of the phrase, measure, beat, or (in some cases) individual note. Some precision is lost in cases where reverb or chord asynchrony makes it difficult to pinpoint the onset, and human error should also be taken into consideration. If pressed, I would estimate the reliability to be no worse than ±20 ms. In general, I like to gather IOI statistics at the level of the tactus, that is, the metrical level that is most salient to the listener (and the performer) and that sometimes provokes toe-tapping. These values can easily be converted to M.M. speeds in beats per minute (bpm).

[7.4] Amplitude is the main physical correlate of intensity, or perceived loudness. The human ear can perceive intensity distinctions across a

trillion-fold spectrum, so a logarithmic measure, the decibel (dB), is used to facilitate comparisons. Regardless of their absolute values, any two sounds at the same pitch level that differ in amplitude by 10 dB (e.g., 0 dB and 10 dB, 90 dB and 100 dB) will have intensities in a 10:1 ratio. The same is not true of linear measures, such as duration, which is one reason why we sometimes use percentages rather than absolute durational values when comparing IOIs. Note that the logarithmic function is curved, so the intensity ratio is not always equal to the difference in dB values. In fact, it is only so in the case of 10 dB differences. However, a useful rule of thumb is that a 3 dB difference roughly indicates a doubling in intensity. [95]

[7.5] It is important to realize that two different scales are in common use among those interested in measuring the loudness of music. One is called "Sound Pressure Level" (SPL), and in this case the baseline (0 dB) represents the normal threshold of hearing (i.e., minimum audible intensity) for pitches around 1000 Hz. This scale is used for comparing sounds "in the air." The other scale, "Electronic Signal Level" (ESL), instead uses the point at which distortion is attenuated as its baseline. The latter scale is used for most forms of electronically mediated sound, including audio editing software. Those who are accustomed to seeing amplitudes in terms of SPL scales, according to which most musical performances fall somewhere between 40 and 90 dB, might be surprised by the readings given by audio editing software, which (if distortion is successfully avoided) fall entirely below 0 dB. Most CD recordings, for example, have ESLs between 90 dB and 0 dB.

[7.6] It should also be noted that, unlike IOIs, absolute intensity values are meaningless in themselves as measures of a performance, because they are dependent upon recording and playback levels. [96] In performance analysis, it is best to concentrate on relative intensities within any given recording. The most useful approach that the current technology enables appears to be the ranking of selected events in terms of loudness. In order to maximize the effectiveness of this, the excerpt should be edited so that the loudest event is amplified to just under 0 dB$_{ESL}$ and all other values are amplified proportionally. This process, which can be carried out automatically using the audio editor's "Normalize" function, widens the range of amplitudes and thus facilitates comparisons. Even if we restrict ourselves to the ranking of relative intensities, caution is required in interpreting amplitude data, since the relationship between amplitude and perceived loudness is extremely complex. Intensity rises in proportion to frequency in the case of isolated pitches, and also in proportion to textural density [97], so it would be simplistic to regard a quantitative change in intensity *as a*

change in loudness. For this reason, it is safest to compare intensity levels in the case of events that are qualitatively similar in terms of parameters other than perceived loudness. Fortunately, significant changes in orchestration occur only three times in the excerpt used in this study (at beats 14.1, 16.2, and 20.2), and the frequency range is relatively small (e.g., the melody lies within the span of a tenth), but in music with a greater variety of orchestration or pitch, intensity values would be less useful measures of loudness, even for the purpose of ranking the dynamic levels of selected beats.

[7.7] Audio editing software can automatically compute the peak sound level (PSL) within any portion of an audio file.[98] First, as I have just explained, the entire excerpt should be edited using the "Normalize" function. Next, the PSL data can be collected for any span of music. Depending on the objective, it might be appropriate to examine each beat (using the onset points located previously as guides), or even each note within passages whose dynamics or accents seem to be of particular interest. Unfortunately, existing technology does not allow us to analyze each voice in a multi-voice texture independently. It is possible to define frequency limits to the analysis (e.g., to analyze the amplitudes for frequencies between 40 Hz and 1000 Hz), but such an approach would discount the contribution of overtones to the perceived intensity of each pitch.

[8] Appendix 3. Quantitative Performance Analyses

[9] Appendix 4. Index of Acronyms

Numbers in parentheses indicate the paragraphs in which terms are introduced.

AM agogic micro-accent (3.3)
DM dynamic micro-accent (3.3)
H_1 first, i.e., most superficial, level of hypermeter (2.1)
IOI inter-onset interval (4.2)
MPR metrical preference rule (2.2)
MS metrical structure (2.1)
MSE metrical structure episode (4.5)
MWFR metrical well-formedness rule (2.2)
PM phenomenal micro-accent (3.3)
PR preference rule (2.2)

Alan Dodson
University of Western Ontario
Faculty of Music
Rm. 210, Talbot College
London, Ontario N6A 3K7
Canada
dadodson@uwo.ca