
Music Expectation by Cognitive Rule-Mapping

EUGENE NARMOUR
University of Pennsylvania

Iterative rules appear everywhere in music cognition, creating strong expectations. Consequently, denial of rule projection becomes an important compositional strategy, generating numerous possibilities for musical affect. Other rules enter the musical aesthetic through reflexive game playing. Still other kinds are completely constructivist in nature and may be uncongenial to cognition, requiring much training to be recognized, if at all.

Cognitive rules are frequently found in contexts of varied repetition (AA), but they are not necessarily bounded by stylistic similarity. Indeed, rules may be especially relevant in the processing of unfamiliar contexts (AB), where only abstract coding is available.

There are many kinds of deduction in music cognition. Typical examples include melodic sequence, partial melodic sequence, and alternating melodic sequence (which produces streaming). These types may coexist in the musical fabric, involving the invocation of both simultaneous and nested rules.

Intervallic expansion and reduction in melody also involve higher-order abstractions. Various mirrored forms in music entail rule-mapping as well, although these may be more difficult to perceive than their analogous visual symmetries. Listeners can likewise deduce additivity and subtractivity at work in harmony, tempo, texture, pace, and dynamics. Rhythmic augmentation and diminution, by contrast, rely on multiplication and division. The examples suggest numerous hypotheses for experimental research.

Received October 13, 1999; accepted for publication January 12, 2000.

MANY psychologists believe that humans apply rules in order to process certain kinds of stimuli. Recent work argues that even infants identify variables, devise symbolic representations, and construct generalizations in order to arrive at contextual linguistic meanings (Marcus, Vijayan, Rao, & Vishton, 1999). Rulelike explanations permeate other scholarly domains as well—for example, the mathematical explanations of visual perspective found in representational paintings (Willats, 1998). Our con-

Address correspondence to Eugene Narmour, Department of Music, University of Pennsylvania, 201 South 34th St., Philadelphia, PA 19081. (e-mail: enarmour@sas.upenn.edu)

tinuing attempts to explain the precise nature of deductive reasoning and its pervasiveness in human life are thus not surprising. One might even say that the desire to discover empirical rules informs all serious intellectual work (Slobodkin, 1992).

The ontological status of cognitive rules has a long and interesting history in psychology (Nisbett, 1993, has a summary). Many have argued that generalized human reasoning relies more on heuristics (choice, decision theory, judgment), category-based similarity, case-based inference, instantiated analogies (whether common or novel), or connectionist networks than on rules. Some antiformalists have denied the existence of abstract cognitive rules altogether. Even those who favor formalism (like Piaget) have contended that rule invocation depends more on natural cognitive development than on conscious learning. Whatever the truth, the evidence for rule-mapping seems to depend on age, domain, level, training, transferability, pragmatic value, and context (Smith, Langston, & Nisbett, 1993).

In music composition, proscription by rule has been de rigueur since the art began. Rules governing counterpoint, orchestration, harmonic doubling, voice leading, and so forth appear everywhere; deconstructing their use is a primary activity of music analysis. Indeed, abstract syntactic rules and the importance of variables permeate music, a fact to which historical documents from early music notation and music theory amply attest.

For theorists of music cognition, however, the interest in rules is not just to demonstrate that the art of composition relies on proscriptions. These theorists want to show instead how cognitive rules channel music listening. Cognitive rules, for example, can underscore the specifics of musical expectation, thereby making available numerous kinds of musical affect. They should be of considerable interest to psychology because they are so highly abstract—more so than in many other systems of communication—and as such, may elucidate how we process the complex temporal events that fundamentally affect our emotions.

Although rules enable mapping to proceed routinely and efficiently in the temporal processing of complex, novel events, music listeners are generally unaware of thinking about cognitive rules or applying them while processing sound. One reason is that abstract cognitive rules, like those found in music, embed deeply in the memory and thus tend to operate unconsciously (Kihlstrom, 1987, shows that much psychological processing is subliminal). There are other reasons as well. In contrast to music, for example, we can visually scan an artwork as long as we wish (Arnheim, 1969). In the visual arts, Solso (1994, p. 147) reminds us that we “think” painting—by which he means contemplate it—even more than we “see” it. Listening to music, however, affords little time for holistic reflection because in the constant stream of stimuli each new item supplants the previous one. Thus ordinary listeners seldom think about music’s effect on the

mind's ear when they experience it. In fact, during music processing it is the composer's rather than the perceiver's use of repeated experience that controls the opportunity for extended reflection. Of course, music often prompts us to think, but because our conscious cognitive resources are limited, too much thinking will suppress one's musical attention and hence detract from the aesthetic pleasure of the experience. To observe that music is replete with cognitive rules is thus not to say that listeners consciously solve problems, verify instantiated variables, or test the pragmatic or empirical value of rule applications. Rather, the invocation of cognitive rules in music is automatic, serving to refine and circumstantiate the listener's expectations.

Different kinds of music often involve the same kind of abstract generalization. One finds the similar iterative applications of additive, subtractive, multiplicative, divisional, and other transformational rules in many different styles. Such abstractions operate across a wide spectrum. Some rules seem "hardwired" and thus perceptual. Others appear learned. Many of the former originate from the bottom up. Others clearly come from the top down. Some are easy to learn. Others require a good deal of training. Some are not congenial to cognitive rule-making at all (e.g., long passages of retrograde).

In this article, we will examine how listeners may apply cognitive rules not only in the usual contexts of augmentation/diminution or in mirror transforms (symmetrical inversion, retrograde) but also in melodic sequences (whether full, partial, or alternating) and in systematic extensions and reductions of melodic implication. Other parameters appear amenable to cognitive rule-making as well (e.g., harmony, texture, tempo, pace, dynamics). As some deduced rules generate strong expectations, along the way we will see how these projections can be denied, delayed, or "overreached" and thus contribute to the aesthetic affect. We will also discuss how listeners may invoke rules purely in retrospect and enter into the musical aesthetic through reflexive game playing. Finally, we will contrast natural rule-making with proscribed constructionism in composition.

Following a case-based, rather than a statistical, methodology, I will present numerous musical instances that are plausibly theorized as rule-governed in terms of expectation mapping. Chosen for their clarity and typicality, the copious examples of music show how applications of rules spread across diverse stylistic populations. In addition, the examples offer cognitive psychologists good models from which to extrapolate experimental materials to test the rule-mapping hypotheses (it is hoped that the mathematical formulations of the musical rules will be useful to those interested in computational modeling). A discussion of the issues and the reservations that one might have about the existence of these rules is delayed until the end of the article, where I will raise experimental questions about the types of rule-mapping seen in the figures.

Similarity and Rule

Iterative cognitive rules combined with formal similarity ($A^0 A^1$) are a stylistic norm of tonal music, powerfully specifying the listener's expectations. Stylistic similarity and the cognitive construction of noniterative rules frequently appear together as well. Because this is so, let us first distinguish similarity from rule so as not to confuse them.

Stylistic similarity involves matching attributes, weighting features, and aligning hierarchical structures (Gjerdingen, 1988; Narmour, 1999).¹ Listeners tend to make content comparisons—and thus establish similarity relations at the manifest level, where specific conditions of similarity—rather than general notions—convey meaning (see Medin, Goldstone, & Gentner, 1993; Pazzani, 1991). Recognition of similarity helps the perceiver construct hypotheses, categorize inputs, and process variables (Medin, Goldstone, & Markman, 1995; Wisniewski & Medin, 1994).

Nevertheless, for all its importance similarity per se does not necessarily preclude rule-mapping. The question for any musical structure, whether formally similar (*AA*) or differentiated (*AB*), is not whether listeners deduce a specifically concrete rule but whether they invoke the same type of abstract rule when variables change in both *AA* and *AB* conditions.²

As we shall see, higher-order abstract generalizations carry over from one form to another. Many different examples will demonstrate that musical rules are not bounded by stylistic similarity (by *AA*) and that rule invocation is possible in formal differentiation (in *AB* without repetition)—for instance, in pitch-differentiated motives with additive rhythms or in passages with successive increments of dissonance. In short, even though in music stylistic repetition and iterative rule appear so interconnected as to be difficult to tease apart, style does not circumscribe rule—and vice versa. Thus one must not equate or confuse stylistic similarity with rule.

Sameness and Style: Exact Repetition

Let us also not confuse stylistic sameness with rule-mapping. Stylistic mapping occurs without invocation of rule, as when one recognizes that an emerging repetition originates from, and thus represents, a generalized

1. Music theorists and psychologists take the concept of similarity for granted, but computational analysis demonstrates how difficult it is to formalize its attributes (see Selfridge-Field, 1998).

2. Cognitive rules governing contexts like these, which often involve formal differentiation (*AB*), are generally more difficult for listeners to construct.

extraopus style structure. Listeners construct style-structural memories from closed forms that composers decide to replicate, either within or without a piece. A passage may initially appear in uniquely differentiated forms of *AB*, but once repeated, the *ABAB* is additionally coded as *AA* on the next level. If repeated enough, the parametric content of such patterns solidify in the memory as schemas, exemplars, prototypes, archetypes, models, categories, and norms.

In online processing, exact repetition of any type, regardless of level, induces a kind of built-in “rehearsal,” which creates the stylistic expectations that arise in working memory. Exact replication, however, reduces temporal uniqueness, assimilates what is materially variable, relegates contrast into mere ornament, and therefore constructs and consolidates style. Purely in terms of stability, musical style is nothing more than exact repetition.³ But for all its obviousness, style is not a trivial aspect of music; indeed, the irresistibility of a given compositional style causes listeners to seek out similar works.

In syntactic systems like music, the top-down schematic mapping of exact repetition does not depend on invoking a simplistic list of features. For exact repetition to be perceived, the match must sufficiently conform to a specifically structured network that organizes clusters of temporal characteristics, features, attributes, and so forth. Thus mapping of exact repetition relies on elaborate frames of reference that match incoming signals to multidimensional properties associated with learned hierarchical structures in long-term memory. In terms of music theory, these specific top-down invocations constitute style structures (Narmour, 1977, 1990, 1992). Such instantiations have a hierarchically specific core, and they map continuously onto incoming information when the alignment of input levels is congruent and when the input is functionally congruent (see the three-dimensional representations in Narmour, 1988 and 1999). Primed with such information in the face of repetition, the listener will then learn to expect a continuation based on prior

3. Of course, creating a compositional style is a cognitively complex activity because composers must consciously decide whether, and the extent to which, replication contributes to the overall aesthetic. But even here the unconscious mind shapes compositional decisions in a nontrivial way. For as a composer writes out a passage and hears it over and over in the mind before committing to a final version, the amount of tolerated repetition may change drastically. This is because a composer cannot objectively determine whether the amount of repetition is redundant and boring or interesting and necessary to the perceptual aesthetic. Composers remain listeners while they create. From a compositional point of view, one cognitive pitfall is to keep repetition at a minimum (as found, e.g., in some serial compositions); another is to have too much of it (as in, e.g., some minimalist or popular styles of music). Great composers are able to balance learned listening and novel stylistic creation within an ever-evolving stance of self-criticism.

experience. Such pattern induction may be shown symbolically as follows (Narmour, 1991):

$$\begin{array}{ccccc} A^0 & + & sA^0 & \rightarrow & A^0 \\ \text{(first} & & \text{(second} & & \text{(second} \\ \text{form)} & & \text{form)} & & \text{form)} \end{array} \quad (1)$$

$$\begin{array}{ccccc} A^0 & + & A^0 & \rightarrow & A^0 \\ \text{(first} & & \text{(second} & & \text{(third} \\ \text{form)} & & \text{form)} & & \text{form)} \end{array} \quad (2)$$

where A^0 = closed form; s = subset; and \rightarrow = implies.⁴ In the first formula, a subset (s) from the second form exactly mimics the first A^0 and thus triggers the expectation of an exact repeat (the second form). The second equation follows logically from this in that once the second form is realized, the expectation of more exact repetition (third form) arises.

Exact repetition—and the listener's inductive mapping of it—is so common in music at every level that numerous musical terms exist to describe it. On lower levels, motivic segmenting, isorhythm, isomelody, ostinato, ground bass, vamp (mechanical meter), and the like are common. On middle levels, we find the terms echo, strophe, canon, voice exchange (*Stimmtausch*), cantus firmus, and talea (longer isorhythmic patterns). Even at the level of entire pieces there exist descriptive terms for exact repetition—for example, da capo forms (like minuet and trio).⁵ Passacaglia and theme and variations, where many parametric relationships remain constant, also involve large amounts of exact repetition.

The phenomenon of repetition has both fascinated and puzzled music theorists. Its obviousness as a compositional device seems incapable of generating aesthetic value, yet its continuing ubiquity inherent in all styles is difficult to explain. Schenker (1954, p. 236) says that “repetition . . . is the basis of music as an art” and that “it creates musical form.” Curiously,

4. The formulation in the present article, based on the implication-realization model, stipulates that repetition plays an integral role in expectation. Meyer (1973, p. 51) says that repetition leads to the expectation of change ($AA \rightarrow B$) instead of repetition ($AA \rightarrow A$) because saturation sets in (Meyer, 1956, p. 152), but this leads to a contradiction because lots of patterns clearly imply the continuation of exact repetition (e.g., a ground bass). Meyer solves the contradiction by invoking context (1973, p. 51). The solution argued in the implication-realization model is that AA implying A is a result of hard-wired, bottom-up processing, whereas AA implying B is a factor of top-down processing, that is to say, of stylistic learning (see Narmour, 1991).

5. To be sure, exact repetition is also implicated in return forms—sonata, rondo, sonata rondo, rounded binary, etc. However, return forms belong to a different discussion. The focus in this article is on contiguous repetitions. I have omitted reference to contrapuntal processes such as fugue, invention, chorale prelude, etc. because contiguous repetitions may or may not follow at the octave. Of course, listeners map exact repetition even in imitation to the extent that parametric relationships remain constant.

some of his followers claim that repetition “tends to create monotony rather than formal interest” (Salzer, 1962, p. 236) and that it is “of no significance in formal analysis” (Green, 1965, p. 80). Others have declared repetition’s meaning to reside in parallel construction and symmetry (Berry, 1966, p. 51). The formulation invoked in the present study is based on the implication-realization model (Narmour, 1991), which insists that because exact repetition causes listeners to map learned expectation, its denial is an essential component of musical affect.

Varied Repetition and Rule-Mapping

In addition to mapping stylistic repetition, listeners encounter incoming material that only partly matches the learned model. When complex input goes against the current mapping of exact repetition, the listener deletes the information irrelevant to the task and begins restructuring his or her expectation (Sadler & Shoben, 1993). Analytically, we must therefore distinguish varied repetition from stylistic replication. The difference between model and varied copy involves an asymmetric judgment because in such contrasts people perceive differentiated features as more heavily weighted than common ones (Goldstone, 1993). This “feature bias” occurs because the structural and functional properties shared between compared forms cause differentiated elements to spring to the fore. Although the closer the relational match the more the common properties will lead to the perception of formal sameness, any variation will tend to appear more salient (Medin et al., 1995). In the symbol string A^0A^1 we see two A s whose features are structurally identical (save for position), but the different superscripts leap even more to the eye. Because in music temporal position does matter—true similarity not being a transitive relationship (Tversky, 1977, p. 329)—we must symbolize dissimilarity. Thus in addition to the symbols of A^0A^0 for exact formal repetition, I shall use the symbols A^0A^1 for varied repetition.

In certain A^0A^1 s the listener may cognitively understand, albeit unconsciously, that the variance itself (the superscript 1) is rule-governed. That is, the varied part of a given repetition may invoke an organizing rule. Specifically, an A^1 following an initial A^0 may cause listeners to comprehend the dissimilarity in terms of a top-down, rule-mapped expectation. This occurs independently of the bottom-up processing taking place and separate from the exact stylistic repetition that occurs. That is, deductive rule-mapping is separate from that part of the first A that the second A replicates and prospectively triggers the inductive non-rule-mapping of the pattern. All other things being equal, the occurrence of $A^0 + A^0$ will gener-

ally imply to the listener another A^0 (Narmour, 1991).⁶ But humans do not just associate varied patterns according to learned gradients of similarity (Pinker, 1999, p. 40) because the variant itself may entail the expectation of more variance according to rule. Symbolologically:

$$A^0 + A^1 \rightarrow A^2 \quad (3)$$

where the A s refer to the exact repetition and the change in superscripts represents the invoked rule. Observe the feedback between the two: repetition constrains the rule, but rule also constrains the repetition (see Jones, 1990).

The Sequence: A Simple Cognitive Rule

A typical example of both deduction and induction is the musical sequence. Sequential transpositions constitute the more popular melodic patterns in tonal music because once listeners deduce the rule they can project the ongoing mapping at new pitch levels and join in the musical play, as it were. In measure 2 of Figure 1, for instance, the zigzagging reversal structure [R⁻] at the quarter-note level occurs four times. Formally, the zigzaggings, the motions from large to small interval, the implied tonic harmony, and the metric-durational patterning produce sameness ($A^0A^0A^0A^0$) and thus strong stylistic induction. But there is a dissimilar property to these forms as well—the sequencing of each motive along a triadic assent ($A^0A^1A^2A^3$; see the reduction underneath). There is also some variation in scale step and interval—the larger ascending skips mixing fifths and sixths, and the smaller descending ones mixing thirds, fourths, and seconds.⁷ (Note also how the length of the ensuing motives increases; the reversals [R⁻] span quarter notes; the differentiated processes [P] take two quarter notes, and the final dyad [2] lasts a whole bar; we shall discuss

6. For this reason, an *AAB* form is effective because *B* is a surprise. “All other things being not equal” is, of course, a style where *AAB* is a norm (as in late eighteenth- and early nineteenth-century music). Because in this invoked style *AA* is heard to imply *B*, rather than *A*, a form of *AAA* can itself be a source of surprise.

7. When discriminating short, relatively contiguous, novel tonal transpositions, there is evidence to show that listeners find contour more easily remembered than exact interval or even scale step (Dowling, 1978, 1991; Dowling & Fujitani, 1971; Dowling, Kwak, & Andrews, 1995). Some say closely related keys enhance this effect (e.g., Bartlett & Dowling, 1980; Watkins, 1985) unless the time span between comparisons is relatively long (Dowling & Bartlett, 1981); in such cases pitch and interval seem to be more important than contour (Edworthy, 1985). Other evidence seems to show that the effects of key distance in identifying sequence are small (e.g., Takeuchi & Hulse, 1992; van Egmond, Povel, & Maris, 1996). Whatever the truth, diatonicity is a powerful organizer of sequential similarity even when exact intervals are not preserved, provided the transposition is not too great (Francès, 1957/1988, van Egmond et al., 1996).

Fig. 1. A simple sequence based on a triadic structure. Beethoven, Sonata op. 13 ("Pathétique"), III, mm. 18–19 (Allegro). R⁻ = prospective reversal of large to small interval; P⁰ = process of similar intervals; (R⁰) = retrospective reversal of similar intervals; 2 = dyad (two-note group); (i) = implicative from the first tone of the dyad.

later how this kind of successive formal lengthening involves rule-mapping.)

Listeners map the exact repetition of this utterly typical passage with little effort, and once the sequential variation gets underway, they also hear that intervallic and scale-step variation follows the implied triadic ascent at the higher level (second staff). Thus, while processing measure 2, there can exist the unconscious invocation of a specific rule as well: "Expect the tones of each repeated reversal structure to transpose up a diatonic third or fourth, consistent with the implied F-minor chord." Of course, to invoke the rule the listener must (1) perceive that the initial subset of the second form (sA^1) is similar and fundamentally analogous to the beginning of the first one (sA^0), (2) abstractly code the variables of A^1 so as to access the relevant rule, and (3) then apply the rule to project the sequential continuation.

Translating the terms into abstract variables, we can formulate a more general syntactic rule for measure 2: "Expect to add an interval of transposition to each tone of each subsequent melodic structure after recognizing the start of the implied repetition." Inasmuch as natural-language rules are cumbersome, we may more economically code this rule in formalized symbols, where A^0 and T represent the input and S and A^1 the output:

$$A^0 + T_n(sA^0) \rightarrow S_n(A^0) = A^1 \quad (4)$$

Specifically, A^0 is the initial form; T the transposition; n the interval of transposition represented in half steps and bounded by a perfect fourth: $\in n\{1,2,3,4,5\}$; small s , the triggering subset; \rightarrow , the implication; capital S , the sequence; and A^1 , the varied form that listeners expect as a result of unconsciously mapping a rule of sequence. The constraint of five half-steps (a perfect fourth) is necessary because experimental data show that distance of transposition is correlated with degree of perceived similarity (van

Egmond et al., 1996; van Egmond & Povel, 1996).⁸ Such rules are probably represented topologically in the memory rather than coded sequentially (Gattis & Holyoak, 1996). Listeners apply them automatically (and probably within milliseconds).

Transposed sequences occur on all levels. They span beats, motives, subphrases, phrases, sections, or even larger formal units. The specific temporal frame constrains the expectation. An abstract rule is, as Jones (1990, p. 194) says, “linked to a listener’s sensitivity to the temporal properties of events.” Experienced listeners can also learn to estimate and thus anticipate the number of times a pattern will sequence because number of repetitions seems to be correlated with length. One finds in the music of Scarlatti, for example, diatonic transpositions at the beat or half-beat level that repeat from 10 to 20 times. In the music of Mozart, sequences that are two beats long rarely repeat more than eight times, whereas sequences that last two to four bars repeat only a few times. Chopin may sequence an entire eight-bar phrase, but usually just once. In the music of Bruckner, which is full of elaborate contrapuntal sequences, phrase-length patterns (four or more bars) rarely sequence more than three or four times. The longest sequence that I know occurs in the development of the first movement of Schumann’s Fourth Symphony, where, with a few harmonic deviations, an entire formal section of 72 bars sequences up a minor third. Although satisfying the analytical definition, it seems highly unlikely, however, that listeners would construe such a lengthy transposition as a mere sequence. Indeed, storing two 72-bar passages ($A^0 A^1$) and then projecting a third one (A^2) according to a derived rule is probably cognitively impossible. Thus the sequential relationship in the Schumann case is purely analytic, a compositional construction.

Listeners can learn to anticipate the pitch distances of transpositions as well. Sequences based on seconds, thirds, and fourths, are common, whereas those built on fifths, sixths and sevenths are not.⁹ Indeed, when we hear a

8. Because in the implication-realization model (Narmour, 1990, 1992) small intervals imply processive continuation and large ones reversal, higher-level sequences built on intervals larger than a perfect fifth will tend to attenuate the expectation of continuation because in terms of the model large intervals at higher levels imply reversal of registral direction. A repetitive context surrounding large, higher-level leaps will nevertheless weaken the reversal implication because formal sameness ($A^0 + A^1$) is not congruent with implications of reversal, which are built on differentiated pitches ($a + b$).

9. Listeners do not normally hear single notes as transpositions, that is, do not conceive of an ascending skip of C-A as a C transposed up a sixth. But we do hear octave transfers (one tone taken up or down an interval of 12 half-steps). Because an identification of octave equivalence stems directly from bottom-up processing, the octave does not ordinarily initiate rule-mapping. For this reason, when we hear motives transposed up or down an octave, we tend to construe them as nearly exact repetitions (perhaps as “echoes”) rather than as sequential implications (indeed, music theorists rarely speak of octave sequences). But if three or more contiguous motivic transpositions by octave occur, one might indeed hear them as sequential realizations, but then only retrospectively.

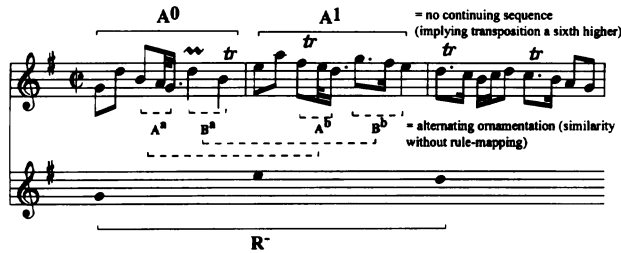


Fig. 2. Ornamented change without rule. Purcell, *Air* (miscellaneous pieces for harpsichord), mm. 1-3. R^- = prospective reversal of large to small interval.

pattern sequenced a sixth away, we tend not to expect another transposition a sixth away. That is because large intervals at higher levels imply reversal, as Figure 2 shows. In other words, although we may perceive a formal transposition (T_n), if n is too large, we will not project a sequential continuation (S_n).

The formal similarity in Figure 2 caused simply by ornamentation—symbolized with superscripted letters as A^a B^a A^b B^b —is not rule-governed.¹⁰ In such contexts of varied repetition, the manifest level is differentiated (ornamented), but the higher-level structure remains the same. As we shall see in Figure 13, however, higher-level changes can be rule-governed ($A^{1,2}$) while lower, ornamented ones nest inside ($A^{a,b}$).

Figure 3 illustrates in Baroque music how pervasive mono-motivic reproduction can be. In measure 1 to the first beat of measure 2 we find much sequential repetition (the structural P^0 -(R^0)s nest within the higher-level (R^0)s, and the P^0 - P^0 s lie within the higher-level (R^+)s). In measure 2 melodic and harmonic sequences join together on each beat, again in linear fashion (see the reductions). In measure 3 the sequential processing appears at the half-bar level, in a contrapuntal texture of voice exchange (zigzagging in both voices). Descending linear patterns (large notes) now occur at the two-beat level (treble: C#-B-A; bass: E-D-C#). The multiple linear sequences on beats (A^0 , 1 , 2 , 3 ; 3 , 4 , 5) thus culminate in a concluding sequence of two beats ($B^{0,1}$). An informal melodic rule would thus be something like, “Expect to add descending diatonic seconds at the beginning of each repetition.” Again in abstract symbols: $A^0 + T_n(sA^0) \rightarrow S_n(A^0) = A^1$,

10. In the case of a tonal sequence at a perfect fifth, the interval sometimes reverses direction or continues to a fourth, in the latter creating a repetition that spans an octave. Once realized, such a sequence sounds more like a return at the octave rather than a continuing sequential implication (see, e.g., the first three bars of the second movement of Haydn's Piano Sonata in Eb, H. XVI:52; the movement in question is actually in E major). According to the implication-realization model, the variable implication of a melodic perfect fifth on whatever level is due to its “threshold property.” Subdividing the octave, the interval can imply either a weak reversal or a weak continuation. For this reason a higher- or lower-level melodic implication of the perfect fifth is context dependent.

motivic forms: A^0 A^1 A^2 A^3
 melodic structures: (R^0) (R^1) etc. p^0 p^1 p^2 p^3

A^3 A^4 A^5 B^0 B^1

(linear) (linear)

Fig. 3. Complex sequences based on the same motive. Bach, Prelude 14, *Well-Tempered Clavier*, 1, mm. 1–3. P^0 = process of similar intervals; (R^0) = retrospective reversal of similar intervals; (R^1) = retrospective reversal of small to large interval.

where A^0 represents the mapping, T the transpositional rule of the subset (s), and S and A^1 the sequential output.

Bach's prelude is suffused with many sequential motives, some toe-tapping 61 of them in fact. Throughout the piece, the procedure of sequence . . . cadence is reproduced numerous times, creating forms of AA . . . B . The sequential repetitions (A^0A^1 . . .) are highly processive and move the music constantly forward. One suspects that many listeners delight in Baroque music because of the mock competitive play between perceiver and composer: "Iterate the sequential rule and accurately predict Bach's musical intentions!"

Sequence and Alternating Segments

One good way for composers to ensure that listeners get aesthetic pleasure is to select only segments of preestablished forms for repetition. These subsets can be subjected to cognitive rule and thus create partial sequences. As in full sequences, the aesthetic of partial sequence relies on denied mapping. That is, its affect depends on the deduction of a transpositional rule ($T_n[sA^0]$) whose implied application is thwarted.

Partial reproduction of exact repetition alternated with partial sequence is found everywhere in music and needs only a little exemplification here. Obeying the earlier transpositional rule ($A^0 + T_n[sA^0]$. . .), Figure 4 begins with a normal sequence on the second beat of measure 1 (A^0A^1). Two new motives (B^0C^0) interrupt the rule-based expectation of a third sequence in measure 2—first a descending process of B-A-G-D (P in the B^0 -part; see

motivic forms: A^0 A^1 $B^0 + C^0$ $B^{0,a} + C^1$

("late")

melodic structures: $(R^0)p^0(R^0)p^0$ $(R^0)p^0(R^0)p^0$ p^0 7 p^0 7 (sequence)

(A^2 if sequenced) (if B^0 repeated exactly) (to C)

Fig. 4. Segmented sequences. Bach, Fugue 15, *Well-Tempered Clavier*, 1, mm. 1–4. P^0 = process of similar intervals; (R^0) = retrospective reversal of similar intervals; P^+ = prospective process of small to large interval; 7 = dyad (two-note group).

“what might have been” in the lower staff) and then an ascending dyadic leap of a D-C seventh (7, the C^0 -part). The first three notes of measure 3 reproduce most of the B^0 of measure 2, but instead of the expected downward leap learned earlier (P^0), we get a linear variant (P^0 ; see the lower staff; the $B^{0,a}$ above the staff symbolizes both the rule-mapping and the ornamental variation). The continuing step to F# is a surprise because it interrupts the exact mapping. Note that the C is not only slightly “late”; the differentiated form ($B^0 + C^0$ following A^0A^1), which ultimately descends to B, leads the listener to dismiss the continuing expectation of ascent (that G-A-B-C will rise to D; see the reduction).

More surprising on the second beat in measure 3 is the leap of the ascending seventh F#-E, which sequences the D-C of measure 2 a major third higher. However, the high E of this sequence is even more unexpected than the analogously high C in measure 2. Of course, the listener does eventually experience a D on the downbeat of measure 4, but that results from a realization of the reversal implication embodied in the F#-E leap.

The rhythmic mapping that “should” follow the leap and the reversal on the D in measure 4—more eighth-notes—is interrupted by a surprising return to the opening sixteenth-note rhythm. In turn, the new melodic motive of measure 4 leads to a partial sequence brought about by the four-sixteenth notes that descend downward (on beat 2, bracketed).

Such are the fascinating alternations of exactly mapped repetition, sequence, and denied sequence in Bach’s fugue subjects.

Partial Sequences of One or Two Tones: Bifurcated Streams

Sequences seem to involve rules that operate above the manifest level. Such high-level structuring of expectations is so common, however, that the theory of their being governed by cognitive rule may appear superflu-

ous. Would not the very frequency of sequences just make them style-structural instances of simple mapping per se? Could not listeners rely just on case-based variation to project sequential continuation? Is it not possible that we simply hear sequences as patterns moving in aural space on some kind of inborn grid such that no rule invocation is necessary? Do we not just reason analogically from one sequence to another?

To answer these questions, consider the zigzagging subject of the Bach invention in Figure 5a. On close inspection, we see here that the applicable rule is not just transposed replication, as it was in Figures 1, 3, and 4. Instead, in Figure 5a, sequence involves rule iteration applied to segments alternating with mappings of exact registral return. Specifically: “Expect after each ascending interval an exact registral return (or retrograde) that

Figure 5 consists of four musical examples, each illustrating a different type of sequence. Example (a) shows a two-part setting of a subject in G major, with annotations for motivic forms (A⁰, A¹, A²) and dyadic structures (3, 5, 8). Example (b) shows a sequence of notes with increasing duration, annotated with A⁰, A¹, A², and A²... adding seconds. Example (c) shows a sequence of notes with increasing duration, annotated with A⁰, A¹, A², and A²... adding a half step to the second and third notes of each grouping. Example (d) shows a sequence of notes with increasing duration, annotated with A⁰, A¹, A¹... adding intervals, B⁰, B¹, B²... sequencing intervals.

Fig. 5. Partial, one-note sequences. (a) Bach, Two-part Invention no. 8, mm. 1–2; aba = exact registral return; (b) Dvořák, Symphony no. 9, IV, mm. 1–5 (Allegro con fuoco); (c) Grieg, Suite from *Peer Gynt*, Anitra's Dance, mm. 15–18 (Tempo di Mazurka); (d) Brahms, Symphony No. 2, I (Allegro non troppo), mm. 17–26.

returns to the initial tone, and transpose every other ascending tone so as to increase the size of the interval in a manner consistent with the chord tones of the implied tonic harmony.”

The rule of sequence, in other words, is partially applied only to ascending tones. Thus, unlike full sequences, partial sequences do not involve higher-level expectations. Instead, the rule governs low-level elements—like the series of items investigated by Restle (1970) and Simon (1972) (for a convenient summary, see Jones, 1974). Nevertheless, in abstract symbols, the rule for partial sequence is the same as full sequence except that the transposed intervals of A^0 involve a subset (s) of the form:

$$A^0 + T_n(s_2A^0) \rightarrow S_n(s_2A^0) = A^1 \quad (5)$$

(where only the second tone [s_2] of the form is transposed).

Because here partial sequence alternates with simple retrograde, a certain amount of repetition is necessary to “prime” the rule-based expectation. The first dyad [3] creates the model. The subset of A^1 that follows—in this case, the exact registral return of the pitch F—plus the transposition of the second subset (the expansion up to C) imply an iterated, sequential expansion on every other tone. The repeated Fs function as an axis of symmetry, while only the second tone of each leap sequences. The initial tone (creating the retrograde), reproduced exactly ($a^0a^0a^0$), consolidates the style, whereas the leaped-to variables (b, c, d) iteratively increase intervallic size while creating formal similarity ($A^0A^1A^2$). When the zigzagging breaks on the high F (m. 2) and rapidly descends the scale, the denial of the iterative alternation gives aesthetic pleasure (the unexpected imitation in the lower voice contributes to the affect; in some sense it realizes the expected return to a low F, albeit an octave below). In short, measure 1 of Bach’s melody makes the prolongation of the tonic chord a highly interesting event.

Music theorists have long recognized alternating patterns of intervallic increment. Toch (1977, p. 125) says that the reiterated tones in patterns like this function as axes, fulcrums, or pivots. Meyer (1989) concentrates on the change in pitch and speaks about the compositional phenomenon of intervallic “stretching.” In psychology the bifurcated streaming that results from such patterns is a well-known aural phenomenon (Bregman, 1990). All such patterns employ a kind of “rubber sheet” technique, not unlike the topological exaggerations found in satirical cartoons (Willats, 1998). The main point, however, is that listeners can cognitively rule-map sequence to partial segments (subsets) of a melody and therefore accurately project variability in pitch, just as they do when confronted with whole-pattern sequencing.

Figure 5b–d illustrates other instances of partial sequence, where various subsets of patterns serve as the iterated variables. Like Figure 5a, measure 4 of Figure 5b sequences only the second note, whereas in Figure 5c

the listener tracks a two-tone segment, which bifurcates down and away from the initial registral axis on B. Here the iterative rule is something like, “Expect to hold the first tone constant while projecting a chromatic sequence on the second and third tones.” Symbolically: $A^0 + T_n(s_{2,3}A^0) \rightarrow S_n(s_{2,3}A^0) = A^1$. Figure 5d, in contrast, shows a minimal one-tone sequence (as in Figure 5a–b) before shifting into regular sequential zigzagging. (Other rules are present in these examples as well—the halving of durations in m. 4 of Figure 5b and the hemiola with complete sequences at the end of Figure 5d; I shall return to these types of rule-governed cognition later.)

As a compositional device, partial sequencing may seem somewhat mechanistic. But because it implies repetition simultaneously with rule application, it sets up the same kind of aesthetic possibilities for interruption and delay, as in any continuation. Consider the partial sequences in Figure 6a–e, from the Quintet of Strauss’s opera *Ariadne auf Naxos*. Here the variables subjected to rule are again single tones. The initial pattern (Figure 6a) involves three intervals (M2, M3, P4), and numerous recurrences of this zigzagging repetition cause the listener to learn to expect three iterated

Figure 6 displays five musical staves (a-e) illustrating partial, one-note sequences varied for aesthetic affect. Each staff shows a sequence of notes with interval labels (A⁰, A¹, A², A³, A⁴, A⁵, A⁶) and interval names (M2, M3, P4, M6, P5, m7) above the notes. The sequences are variations of a three-note motif. The lyrics are: a) Es gilt, ob Tan - zen, ob Sin - gen tau - ge; b) Es gilt, ob Tan - zen, ob Sin - gen tau - - - - ge; c) Es gilt, ob Tan - zen, ob Sin - gen tau - ge; d) Doch wie wir tan - zen, wie wir sin - - - gen; e) Doch wie wir tan - zen, wie wir sin - - - - - gen.

Fig. 6. Partial, one-note sequences varied for aesthetic affect. Strauss, five related motives from *Ariadne auf Naxos*, act 2, between rehearsal letters 79–100 (Andante mosso).

increments in the melodies. Hence stylistic repetition suppresses any expectation that the intervals will expand further.

However, after several repetitions, Strauss takes the Leitmotif further, to the interval of a fifth (Figure 6b). He then expands the terminal melodic interval in surprising ways, to a sixth (Figure 6c), a seventh (Figure 6d), and, finally, an octave (Figure 6e). Accordingly, during the Quintet listeners revise their originally learned expectations (that the melody will span only a third). The rule learned for Figure 6a–c would be something like, “Expect the terminal note of each form (A) to sequence upward by step or third with the initial tone returning to the starting register at each repetition.” But for Figure 6d–e the revised rule would be, “Expect the terminal note of each form (A) to sequence triadically once the stepwise sequence is broken (while continuing to expect return to the initial register).” Thus what could be a mechanistic compositional device becomes in Strauss’s masterful hands a source of developing affective delight. (To aid in learning the style of the melody, the Prologue, actually written after the opera, introduces this Leitmotif to the audience no less than 13 times, although not in the intervallically systematic way heard in the Quintet.)

Rachmaninoff’s cantata, *The Bells*, is another work that makes extensive use of one-note sequencing (here the rule governs descending intervals). The passage first appears in the coda to the first movement (Figure 7a), where partial sequencing occurs before a full sequence moves the pattern down by structural thirds (with some intervallic change toward the end; see the reduction). Hence the listener hears the partial sequence nested within the full sequence. Figure 7b illustrates the transformation of the same Leitmotif in the second movement. Varied repetition (mm. 1–5), diminution (m. 6), and successive augmentations (mm. 8 and 9) prepare the advent of the theme proper (Figure 7c). The theme, which permeates the whole third movement, is sequenced upward, twice (in mm. 18 and 22; in the middle of the third movement the Leitmotif returns in another guise, as an ostinato).¹¹

Figure 8 requires listeners to invoke the same rule except that here, following an exact repetition, the alternating sequential operation applies to different registral directions so that two diverging streams emerge. Constructing the rule is made cognitively easier by the clearly implied harmonic context: “Following the mapping of exact intervallic repetition (A^0A^0), expect bifurcated successions of sequential lines that alternate stepwise motion in contrasting registral directions, iterated according to the voice leading of the implied chromatic harmony.” As before, near registral return ($aba^1b^1 \dots$) underscores the melody of two separate lines. Note that the

11. The partial sequences of Figure 7 and the theme itself bear a strong affinity to the *Dies Irae* melody, which figures strongly in the last movement’s preoccupation with death.

Figure 7 displays three musical examples (a, b, c) illustrating nested sequences in Rachmaninov's *The Bells*. The notation includes various annotations such as A^0 , A^1 , A^2 , A^3 , etc. (expanding intervals), ff , p , mf , f , $dim.$, $etc.$, p , $dim.$, $Augmentation$, $Further augmentation$, $(same procedure: A + A + A \dots)$, $(whole pattern sequenced by fourth)$, $p dolce$, $poco cresc.$, and $(sequenced by third)$.

Fig. 7. Partial sequences nested in regular sequences. Rachmaninov, *The Bells*; (a) rehearsal letter 25 (coda), I (Meno mosso, Maestoso); (b) *ibid.*, II, mm. 1–10 (Lento); (c) *ibid.*, II, mm. 13–24 (Lento).

octave that ends the sequencing appears in rhythmic augmentation (we will later see a similar closural change in m. 3 of Figure 10a and in m. 4 of the antecedent phrase of Figure 13).

Returning to the questions posed at the beginning of this section: Despite their very considerable stylistic contrasts, Figures 5–8 are all alike in terms of rule (so are Figures 9–12). Learned mapping alone does not adequately account for this fact. Although it is plausible to imagine a complete sequence as an example of case-based expectation when the variation remains largely the same as the model, it is difficult to believe that simple

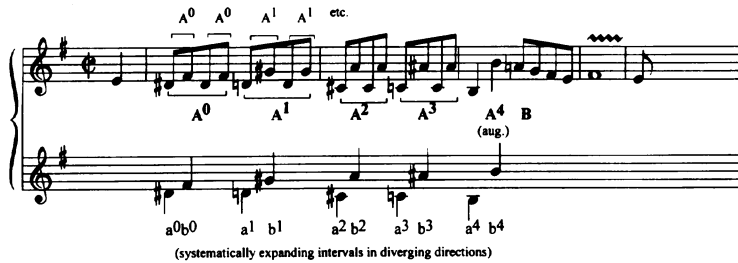


Fig. 8. Alternating one-note sequences. Bach, Organ fugue in E minor, BWV 458, mm. 1–4.

mapping constitutes a sufficient explanation when variation selectively expands certain intervals by an operation of addition that was not present in the initial model. As we have seen, the higher-order rules of partial sequence carry over from one instance to the next. Once we abstract the rule of partial sequence, we are able to systematize many stylistically different pieces efficiently under one rubric. Where possible, our cognitive mechanisms appear to codify diverse expectations at some very high level according to a single abstract rule.

Alternating Sequential Applications

Just as listeners are able to use abstract rules to map formal variation accurately, so they can also variously apply rules when discontinuously alternating sequences occur. In the solo violin sonata of Figure 9a, one melody separates into two, the top one sequencing by ascending thirds, the bottom voice descending by the same interval. Theorists call melodies like this “compound,” referring to the fact that both composers and listeners conceive such passages as independent (though related) melodies. One rule variously governs both lines (formal alternation accompanies the expected repetitions; As are followed by Bs). With the repeat of the successively rising and falling patterns, the listener confidently predicts the separate continuations according to the following rule: “Expect bifurcated lines in contrasting registral directions with separate sequences that mirror each other.” When the last rising line continues the scale upward (m. 4) and breaks the stylistic sequential mapping, the listener experiences an exhilarating change (C).

Alternating rule applications can apply to partial segments as well. In Figure 9b, listeners anticipate that each of the bifurcated lines follows a sequential rule and hence alternates the stylistic mapping. The melodic line of the main motive—doubled in the alto and the bass—descends in a quasi-sequence (As) accompanied by a sequential melodic rise (Bs). Together,

(sequential expansion by divergent thirds)

(diverging harmonic intervals)

gm: i III ii V
(quasi melodic sequence)

(increasingly larger melodic skips with diverging pitches)

Fig. 9. Alternating full sequences. (a) Bach, Violin sonata in B minor, 9 (Double), mm. 14–18; (b) Rachmaninov, Prelude op. 23 no. 5, mm. 12–14 (*Alla marcia*); (c) *ibid.*, mm. 78–80.

these contrasting sequences—one linear, the other triadic—create a strong contrary motion. The listener constructs a stylistic rule-map that predicts the widening harmonic intervals. For each stepwise descent in the main melody, the continuously rising triadic line in the accompaniment ensures the perception of an organized zigzagging (until the cadence), in spite of the nonsequential harmony (i-III-ii-V).

In Figure 9c, which is toward the end of the same piece, the large jumps of the descending motives (Cs) this time produce a compound melody. Although we find in the stylistic repetition bifurcated contrapuntal lines, contrary voice leading is not the whole explanation. The vaulting tones of the main motive follow their own rule-governed partial sequence. The soprano ascends by step ($A\flat$ - A - $B\flat$) until the pattern is broken at the cadence, while those of the bass descend in accord with the harmonic progression. The listener cognitively deduces that the intervals between the two lines of the similar motives involve melodic rules that stipulate a growing space between the initial and terminal tones (the accompaniment this time—the *B*-part—is not part of the sequence).

Cognitive Rule and Art: Denied, Delayed, and Overreached Intervallic Expansion

We tend to think of rule-mapping and aesthetic production as contradictory outputs, the one precluding the other. Rule is common and matter of fact, art is individualistic and fanciful. In music, however, rule and expectation are frequently correlated. Thus denial or delay of a projected rule can produce a strong affect. The implied intervals in the undulating patterns of Figures 6 and 9a–c, seen earlier, systematically increased in size. The formal similarity made it easy for listeners to apply relatively simple cognitive rules of expanded intervallic motion in order to project the continuation accurately. Partial sequences, however, can result in partial realizations. A delay of an expected interval along with formal change can occur while the zigzagging registral return (*aba*) continues. Partial realization may overreach the implied intervallic expansion. Both delay and overreach may occur together. All contribute to the melodic aesthetic.

Consider the famous oboe solo of Figure 10a from the second movement of Brahms's First Symphony. Here the pattern to be repeated begins on beat 3 of measure 2. Each undulation expands the initial interval of the motive while returning to the starting registral pitch ($C\sharp$) and holding fast to the terminal segment of $B\sharp$ - $D\sharp$. Thus $C\sharp$ - $D\sharp$ (the $M2$ of A^0) is replaced by $C\sharp$ - E (the $m3$ of A^1) which itself is replaced by $C\sharp$ - $G\sharp$ (the $P5$ of A^2). Thus we find a rule of intervallic expansion via partial sequence within a triadic framework (recall also Figure 5a). The two eighth notes of the octave on beat 3 of measure 3 break the repetition, but the listener's rule-governed expectation is partly realized because, had the sixteenth-notes continued, the high $C\sharp$ would have followed anyway (see the reduction). Hence the esthetic effect of the high $C\sharp$, occurring one sixteenth "too late," depends on *B*'s denial of the formal implication ($A^0A^1A^2 \rightarrow A^3$), which momentarily

Figure 10 consists of two musical examples, (a) and (b), illustrating alternating partial sequences.

(a) Brahms, Symphony No. 1, II (Andante sostenuto), mm. 39–43. The notation shows a melodic line in treble clef with a key signature of three sharps (F#, C#, G#) and a 3/4 time signature. The tempo is *Andante sostenuto*. The music is marked *p* (piano) and *dolce*. Above the staff, intervals are labeled: A^0 (M2), A^1 (m3), A^2 (P5), and B^1 (slightly late). Below the staff, the corresponding pitch classes are listed: $a^0 b^0 a^1 b^0 a^0 b^0 a^1 b^0 a^0 c^1 a^1 b^0 a^0$. A note below the staff indicates "(intervals increasing in size)".

(b) Brahms, Symphony No. 1, I (Un poco sostenuto), mm. 29–32. The notation shows a melodic line in treble clef with a key signature of two flats (Bb, Eb) and a 3/4 time signature. The tempo is *Un poco sostenuto*. The music is marked *p* (piano) and *espr.* (espressivo). Above the staff, intervals are labeled: A^0 , A^1 , A^2 , B^0 , and B^1 . Below the staff, the corresponding pitch classes are listed: $a^0 b^0 a^1 b^0 a^0 c^1 a^1 b^0 a^0$. A note below the staff indicates "(expanding intervals)".

Below the musical notation, there is a section titled "registral dir. (up/down):" with a sequence of notes: A, A, A, A, A, A... Below this, there is a section titled "interval:" with a sequence of notes: A^0 , A^1 , A^2 , B^0 , B^1 . Below this, there is a section titled "harmony (cm):" with a sequence of notes: V, a^0 , a^1 , a^2 , a^3 , a^4 . Below this, there is a section titled "registral return:" with a sequence of notes: a^0 , b^0 , a^1 , b^1 , a^2 , b^2 , a^3 , b^3 . A note below the staff indicates "(overreached)".

Fig. 10. Alternating partial sequences. (a) Brahms, Symphony No. 1, II (Andante sostenuto), mm. 39–43; (b) *ibid.*, I, mm. 29–32 (Un poco sostenuto).

interrupts the partial sequencing before realizing the triadic expectation. (One notes a similar use of the durationally cumulative octave in Figure 8.)

The no-less-famous oboe solo from the introduction of the first movement of the same piece follows a similar procedure (Figure 10b), although here a subtle division between the expected ongoing registral change and the expected intervallic expansion shapes the denial. The zigzagging continues throughout the phrase ($A^0 A^1 A^2$, bottom staff), but because the chord moves to the tonic on the downbeat of measure 2, the listener projects a new partial sequence, one that involves triadic expansion ($B^0 B^1$, bottom staff). This continues as expected until the stylistic break on the low D, the last eighth note of measure 2, where two octaves now ensue to create a full sequence ($C^0 C^1$, upper staff). The rule making, however, does not end completely with this sequential change because of the constant eighth notes and the predictable registral zigzagging. Hence listeners hear the mildly dissonant high D (downbeat, m. 3) as a slight overreach, a “wrong-pitch” realization of intervallic expansion despite the full sequence established by the octaves. In short, the high, dissonant D’s aesthetic affect results from both the continuing registral zigzagging and the overreached realization.

Continued formal similarity is not the only context where listeners feel the momentary denial of rule-governed increment. Indeed, we find delayed

realizations throughout tonal style in many kinds of formally differentiated patterns. Consider Figure 11a, a typical case where realized expansion with a specifically implied triadic increment occurs as expected (bottom staff). A partial sequence expands the beginning interval of the initial motives, while the last four notes of the A^0A^1 (A-G-F#-E) remain constant. As the harmony changes, the syncopated third motive (A^2 -B) breaks off the exact repetition. But a C surprisingly takes the leap up to a high G (m. 3), hence renewing the intervallic expansion that the formal differentiation was about to make the listener forget. The rapid sixteenthths of *D* differentiate the rhythm further. Nevertheless, the unexpected, almost two-octave leap to the high D (*E*) mimics the rule-governed style established earlier. What we see is that rules governing partial sequence are strong enough to be cognitively relevant even when formal differentiation (*C*, *D*, *E*) and a considerably delayed realization take place (note also the successive halving of motivic length in m. 3; I shall return to this procedure later).

Although much more dramatic, Figure 11b is nonetheless highly similar in procedure to Figure 11a; indeed, it also trades on the same techniques of delay and overreaching seen earlier in Figure 10b (Mahler learned more from Brahms than scholars generally acknowledge). The initial motive employs a typical melodic strategy of romantic composition where the affect of the denial is maintained throughout: Imply a linear ascent but then

Figure 11 consists of two musical examples, (a) and (b), illustrating alternating partial sequences with aesthetic denial. Example (a) is from Haydn's Sonata for piano, H.XVI:39, III, mm. 31–36 (Prestissimo). It shows a melodic line on a treble clef staff and a harmonic line on a bass clef staff. The melodic line is divided into segments labeled A^0 , A^1 , A^2 -B-C-D, and E. The harmonic line shows expanding intervals and a delayed G. Example (b) is from Mahler's Symphony no. 4, III, mm. 205–210 (Leidenschaftlich und etwas drängend). It shows a melodic line on a treble clef staff and a harmonic line on a bass clef staff. The melodic line is divided into segments labeled A^0 , A^1 , A^2 , $\frac{1}{2}A^2$, $A^{3,a}$, and $A^{4,b}$. The harmonic line shows expanding intervals. Both examples include annotations such as '(expanding intervals...)' and '(G delayed)'.

Fig. 11. Alternating partial sequences with aesthetic denial. (a) Haydn, Sonata for piano, H.XVI:39, III, mm. 31–36 (Prestissimo); (b) Mahler, Symphony no. 4, III, mm. 205–210 (Leidenschaftlich und etwas drängend).

skip to a dissonant tone such that its implied resolution in the opposite registral direction strengthens the expectation of a delayed realization. Following this strategy, Mahler exactly repeats the first three tones of the *A*-motive while sequencing its last three tones. This increases the size of the leap from a minor third to a tritone (with a perfect-fifth grace note interposed), and it sets up the rule-governed expectation that the established intervallic expansion will continue (as it does). A segment of the reversing leap in the third motive (one half of A^2) repeats, delaying the projected increment, which nevertheless eventually takes place (m. 5). But unlike the octave overreach seen earlier in Figure 10b, here the vaulting leap is extravagant, a dramatic fifteenth (octave plus a major seventh). The stretch goes way beyond that projected by the rule while augmenting the rhythm (and thus symbolized as $A^{3,a}$). Another variation brings the pent-up passage to a further augmented climax ($A^{4,b}$).¹²

Of course, partial sequence can be denied altogether. For just as two ascending melodic pitches like C-D ($a^0 + a^1$) may not continue upward as expected, so implied continuations based on $A^0 A^1$ forms of partial sequence may not be realized either (Narmour, 1996). We have already seen this in connection with the complete sequences of Figure 4, where a *B* form interrupted two *As*. Figure 12a illustrates the possibility of complete denial in the partial sequence, where on the basis of $A^0 A^1$ the listener formulates a rule of partial sequence and projects a third form only for it not to be realized. In measure 3 Franck sequences the opening skip of the fourth (F-B \flat) to a skip of a sixth (F-D \flat), but then with the return of the F-B \flat he denies that expectation (that the leap will expand triadically to an octave; see the reduction). Such denial of stylistic expectation—where only model (A^0) and varied repetition (A^1) occur—is utterly common.

Interestingly enough, in the last movement of the same symphony (Figure 12b), Franck again expands the fourth to a sixth and then interrupts it (see mm. 5–6 and 7–8). However, unlike the example from the second movement (12a), here motivic realization of the implied expansion eventually occurs in the course of developing the theme (mm. 13–19): The fourth moves to a perfect fifth, a major sixth, and finally, via motivic segmentation, to an unanticipated and somewhat surprising octave (the listener expects a minor seventh). The descending eighth-note filling-in of the intervals, mimicking measures 5–6 of Figure 12a, thus not only transforms the

12. Figure 11b is Mahler's most dramatic development of this Leitmotive, which pervades the entire movement. The first time it appears (mm. 67–70), segmented repetition stops the expansion. The second time (mm. 84–90), the large and surprising skip at the end of the melody occurs, but in a lower register. The fourth time (mm. 298–304), the motive appears pianissimo with some augmentation and thus does not climax. The fifth and last time (mm. 326–338), the huge skip is rhythmically augmented and consonant.

Figure 12 consists of two musical excerpts, (a) and (b), from Franck's Symphony in D minor. Excerpt (a) is from the second movement, measures 17-24 (Allegretto). It features a melodic line in the upper staff and a bass line in the lower staff. The melodic line is marked *p* and *cantabile*. It contains four phrases labeled A⁰, A¹, A⁰...B, and C. The A⁰ and A¹ phrases are connected by a slur. The A⁰...B phrase is marked with an exclamation mark and the text "Bb unexpected". The C phrase is marked with a triangle. The bass line is marked with a triangle and the text "(expanding intervals)" and "F implied". Excerpt (b) is from the third movement, measures 7-26. It features a melodic line in the upper staff and a bass line in the lower staff. The melodic line is marked *dolce cantabile* and *p*. It contains two phrases labeled A⁰ and A¹. The A⁰ phrase is marked with a triangle and the text "(fourth expanded to a sixth)". The A¹ phrase is marked with a triangle and the text "(fourth... expanded to a fifth)". The bass line is marked with a triangle and the text "cresc." and "dim.". The score is in D minor and 3/4 time.

Fig. 12. Denied and then realized partial sequence. Franck, Symphony in D minor, II, mm. 17–24 (Allegretto); (b) *ibid*, III, mm. 7–26.

theme but also links both movements together and prepares a reprise of the second movement's theme in the final movement.

Obviously, tonal composers exploit both realizations and denials of partial sequence. The fact that both procedures occur everywhere in tonal style is a function of their being firmly grounded in the rule-mapping processes of human cognition.

Form, Cadential Closure, and the Sequence

Sequences produce a cognitive and aesthetic effect on music listening that the word “transposition,” which only describes the compositional operation, does not account for. Melodic sequences, for example, have implicative, formal, and rule-governed lives that can operate independently of, and thus lie noncongruent against, harmonic motion. For this reason, in music analysis melodic sequences cannot be reduced to harmonic progressions. Indeed, it is often the noncongruence between melody and harmony that engages listeners—as if two independent activities were present. Melodic activity involves formal repetition, implicative projection, and rule; in contrast, harmonic activity is through-composed, ongoing, and processive.

Fig. 13. Effect of sequence on closure. Mozart, Sonata K. 281, III, mm. 1–8 (Allegro). 2 = dyad (two-note group).

In this connection, consider Figure 13, where linear sequences appear in measures 1–2 (two motives, four quarter notes long with the upbeats), measure 3 (descending melodic dyads [2] on each beat), and measures 4–6, the start of the consequent phrase (third beat, each motive again four quarter notes in length, though varied this time). In measures 1–2 of the antecedent phrase, the overall progression is V/ii–ii–V–I and thus harmonically closed. But every exactly reproduced parameter—melodic contour, voice leading, durational patterning, meter, register, dynamic, and texture—contributes toward formal similarity and thus toward the expectation of another repetition ($A^0 + A^0 \rightarrow A^0$). Those motivic aspects that imply another A^0 therefore weaken the closure of the V–I progression.

Closural attenuation by formal repetition is obviously appropriate for any opening gesture.¹³ In contrast, the differentiated B^1 of the A^0A^1B form strengthens the closure of the half-cadence of the antecedent phrase in measure 4. By the same token, the nested, lower-level sequential dyads [2s] in measure 3 to the cadence in measure 4 themselves create a little form of $A^0A^1 \dots B$. This too reinforces the cadential closure on the dominant in measure 4.¹⁴

13. It is also possible, however, to hear the D–C quarter-notes of the half-cadence itself (downbeat, m. 4) as an augmentation (aug.) continuing the dyadic sequences and thus partly continuing the formal similarity.

14. A difficult task for composers is to maintain a level of nonclosure sufficient to hold the interest of the listener. Toward that end, using sequential repetition, which implies continuation, is an effective compositional device even when harmonic closure occurs.

In the consequent phrase, measures 4–5 are analogous to measures 1–2 except that ornamentation weakens the low-level formal similarity. The rule-governed linear sequence (A^0A^1) is still present at the higher level, but on the manifest level, the first motive of the original melody now appears in a more animated figuration. I symbolize the non-rule-governed change of content in the context of varied repetition as A^aA^b , but because both ornamentation and rule-mapping (sequence) are present here, the initial form in the consequent phrase becomes $A^{0,a}A^{1,b}$. (The A^0 tells us that this form has appeared before, the A^1 that there is a sequence present, and the A^a and A^b that variation occurs in both A^0 and A^1 ; recall also the discussion of Figure 2).

As to the *B*-part in the consequent phrase (mm. 4–8), if the theory of form, sequence, and closure argued here is correct, then the cadential closure (mm. 7–8) of the more formally differentiated A^2A^3C , which imitates the sequence found in the antecedent (cf. mm. 1–2 to 5–6), is even greater. Measure 6 displays much less surface similarity than measure 3, and that is reflected in the higher levels (contrast the up-down leaps of *C* with the harmonic tenths descending by step in *B*; see the reduction).¹⁵ But the music is more subtle than this analytical comparison suggests. For the lower level, descending triadic pattern of measure 7 (C^1 , between the staves) is a symmetrical quasi-retrograde of C^0 , which slightly weakens the formal differentiation of C^0 at the higher, two-bar level. Moreover, measure 6 creates an important aesthetic moment in the phrase by triadically enveloping the analogous F-Bb leap in the antecedent phrase (cf. m. 6 with m. 2).¹⁶ So despite the formal differences between the end of the antecedent and the end of the consequent, these subtle similarities somewhat mitigate the clausal strengthening of the differentiated A^2A^3C .¹⁷

This subtlety reminds us how important it is to distinguish the purely formal aspect of A^0A^0 implying A^0 (exact repetition) from the actual content of the implied variance (superscripted). If listeners prospectively hear any A^0A^1 variance as rule governed, they will expect continued rule application (not only where $^0 + ^1$ implies—and is realized—by 2 , as in Figures 1 and 3, but also where $^0 + ^1$ implies a continuation that goes unrealized, as in Figure 13).

15. I shall pass over the fact that both *B* (encompassing the half-cadence) and *C* (spanning the full cadence) are hierarchically “deeper” than the *As*. This too increases the formal differentiation and thus the cadential closure in both the antecedent and the consequent.

16. One might think that the trill to the cadence in measure 8 is also a microscopic *AA* . . . *B* form (up-down-up-down . . . followed by the quarter-note Bb). However, the very fast pace of the trill prevents the listener from parsing each undulation on the thirty-second-note level as a tiny structure.

17. Forms like these tend to be cumulative in temporal design ($1 + 1 + 2$) and thus in terms of time span also closed. They are so common in tonal music that listeners can learn from an initial *AA* to expect a *B* that is a temporal multiple of *A*.

Not all variance, of course, generates prospective rule-mapping. In measures 6–7, the listener's ascription of the rule-governed retrograde is “backward looking,” purely retrospective because in perceiving mirror forms one has to wait until the event concludes before deducing a post hoc rule that connects copy to model. I shall return to the topic of retrograde below.

There is a third kind of variance that is neither prospectively or retrospectively rule governed. As we have seen, ornamentation often occurs in contexts of repetition.¹⁸ Thus we must discriminate this kind of change in content from rule-governed change. As discussed earlier, in order to track the former I use superscripted letters ($A^a + A^b + A^c$). The varied ornamentation in the alternating forms of Figure 2 ($A^a B^a A^b B^b$) illustrates this, and Figure 13 shows that both non-rule-governed ornamentation and rule-governed sequence can occur together when the former nests inside of the latter ($A^{0,a} A^{1,b}$).¹⁹

In terms of processing repetition, we find therefore (1) the exact mapping of previously learned style structures or $A^0 A^0$; (2) the mapping of inferred rules along with some degree of exact repetition or $A^0 A^1$ (whether applied prospectively to continuations, such as sequence, or retrospectively to mirror forms); and (3) ornamental change with no rule-mapping or $A^a A^b$. Table 1 shows exact and varied repetition across a spectrum from sameness to similarity to difference along with a summary of each type's characteristics.²⁰

The cognitive impact of musical form—whether motives are exact ($A^0 A^0$), similar ($A^0 A^1$, $A^a A^b$), or different (AB)—is thus not totally determined by functional content, as commonly believed (see, e.g., Randel, 1986). Form is not just a convenient analytical summary; it is a crucial psychological property of expectation and thus the ongoing musical aesthetic.

Augmentation and Diminution: Noniterative Rules Governing Duration as a Variable

Cognitive rules also specify the mapping of augmentation and diminution. The former involves multiplication; the latter, division. Both preserve durational proportions and thus constitute noniterative rules. To rule-map either procedure onto one's expectations, one first comprehends that dura-

18. Ornamental change, however, is rarely arbitrary. It too relies on cognitive explanation, but space does not permit me to explore the issue of form and content here.

19. Note that motivic differentiation occurs on the manifest lower level (hence the superscripted letters), whereas formal similarity (As , and Bs alternating) points to the higher level (without which we could not categorize the phenomenon of ornamentation).

20. Rule-mapping (A^1) and mere content change (A^b) can occur simultaneously (recall Figures 2 and 11b, and see Figures 15a, 16c, and 25b).

TABLE 1
Intraopus Forms

Form	A ⁰ A ⁰ sameness (reproduction)	A ⁰ A ¹ similarity (variation)	A ⁰ A ^b dissimilarity (ornamental variation)	AB difference (contrast)
Level properties between forms	Both forms exactly alike	Both somewhat alike	Only higher levels similar (manifest levels different)	All levels different
Mapping properties	Stylistic mapping only	Rule-mapping plus stylistic mapping	Weak mapping on higher level; no rule-mapping	No stylistic mapping; rule-mapping possible
Cognitive orientation	Prospective	Prospective	Initially retrospective	Initially retrospective
Higher-level implicative strength	Strong	Strong	Weaker	Weakest
Cognitive level of attention	Higher level (can recede to ground)	Higher level (focus on process)	Lower level (site of ornamentation)	Lower level (site of differentiation)

tion alone is the parameter on which the rule operates and then projects the temporal change of all subsequent mappings of the learned pattern. Because original proportions remain the same in both augmentation and diminution, melody is mapped simply, as exact repetition. Indeed, melodic sameness facilitates perceiving the relationship between the durational model and its transform. We are less aware of augmentation or diminution if the melodic intervals are differentiated.

The prevalence of these two techniques at every level and dimension demonstrates that rule making is cognitively operative in parameters other than melody. Augmentation and diminution also show that cognitive rule-mapping in music is not dependent on formal similarity (AA) inasmuch as these operations do not require the occurrence of exact formal repetition.

We have already witnessed augmentation in Figures 7b, 8, 11b, and 13 and diminution in Figures 5b and 7b. As analysts have written extensively about both, particularly with respect to contrapuntal music, the phenomena need no extended discussion here. In order to project the continuation, listeners rule-map a one-time proportional increase in durational value in augmentation or a one-time proportional decrease in diminution.

In Figure 14a, an often cited case, augmentation of the fugue subject appears in the salient top voice and continues for five bars. In the bass and then in the alto, the contrapuntal voices occur in their (nearly) original durations. In the augmentation, the listener hears each duration of the original as multiplied by two. Single eighth notes in the original become quarter notes in the augmentation (two eighths tied). For example, the tied quarter note A# in the bass of the first measure is three eighths in length (the original duration), so the high A# in the soprano of measure 2 is six eighths (i.e., three quarter notes). Although texturally dense stretto passages like this put a heavy processing load on the listener, cognitively untangling the rule interplay in the counterpoint greatly contributes to the aesthetic experience.

Figure 14b, another famous case, illustrates diminution and augmentation at the same time. The original subject is shown in small notation above

Figure 14 consists of two parts, (a) and (b), illustrating musical techniques of augmentation and diminution.

Part (a) shows a musical score for Bach's Fugue 8, Well-Tempered Clavier, measures 77-84. The score is in G major (one sharp) and 3/4 time. It features three staves: Soprano, Alto, and Bass. The Soprano staff shows the 'Augmentation' of the fugue subject, where durations are multiplied by two. The Alto and Bass staves show the 'Original' subject. The Soprano staff also includes a section labeled '(Stretto)' and another labeled 'Augmented and varied'. The Alto staff is labeled 'Original (with slight variation)'. The Bass staff is labeled 'Original'.

Part (b) shows a musical score for Beethoven's Sonata op. 110, IV, measures 1-4 (fugue subject) and mm. 127-130 (Allegro, ma non troppo). The score is in D major (two sharps) and 3/4 time. It features two staves: Treble and Bass. The Treble staff shows the 'Fugue Subject' in small notation, followed by 'Augmentation (multiply by 2)'. The Bass staff shows the 'Diminution (divide by 3)'. The Treble staff also includes a section labeled 'etc.'.

Fig. 14. Augmentation and diminution. (a) Bach, Fugue 8, *Well-Tempered Clavier*, 1, mm. 77-84 (in open score); (b) Beethoven, Sonata op. 110, IV, mm. 1-4 (fugue subject) and mm. 127-130 (Allegro, ma non troppo).

the Figure, which, having previously heard it some 16 times, the listener has in memory. Thus one can readily deduce both augmentation and diminution and, if properly performed, perceive the proportional conservation while mapping the dense counterpoint. Discounting the tied upbeat, we hear the durations in the top voice as double that of the initial ones. At the same time, the zigzagging diminution in the motives in the imitative bottom voices divides the fugue subject's duration by three (dotted eighth notes = $3/8 \div 3 = 1/8$). (A further diminution, where the initial notes of the subject take only one-and-a-half beats, occurs a few bars later and employs inversion as well.)

Symbolically, diminution and augmentation are the simplest cognitive rules. Where d is the duration of a note, v is its value (given $v \in A^0$), q is the quantity by which it is divided or multiplied, and v^1 is $v^1 \in A^1$, the rule for diminution is:

$$d(v) \div q = d(v^1) \quad (6)$$

For augmentation, the rule is:

$$d(v) \times q = d(v^1) \quad (7)$$

Listeners can recognize augmentation or diminution on very low levels, even within a single motive. We saw a very brief augmentation in Figure 13 (m. 4); Figure 15 shows two other cases. Because of its brevity, the diminution in Figure 15a (G#-A#-B)—also a sequence—allows the listener no opportunity to map the relationship prospectively. The delight in deducing that the durations of the initial motive have been divided by two is thus at first purely retrospective and formalistic. Similarity relations, however, may be asymmetrically directional (Tversky, 1977), depending on whether the relationship of model to variation is one of greater to lesser or lesser to greater (Medin et al., 1995). That is, it may be that the relatively faster pace of the second motive reduces the amount of similarity perceived, which would weaken the listener's sense of diminution. Further ornamentation disguises the beginning of the second sequence ($A^{2,a}$); so the listener's understanding of it is quite retrospective as well. However, whatever asymmetry is present may be more than offset by the sequential similarity ($A^1 A^2$). Be that as it may, A^1 of Figure 15a actually involves two rules—diminution (rule 6) embedded in complete sequence (rule 4). This may be expressed as:

$$A^0 + T_n(A^0 [d(v)/q = d(v^1)]) = A^1 \quad (8)$$

By contrast, Figure 15b moves from lesser to greater, as it were. Having heard the repetition of a terminal motivic segment and then its partial sequence ($A^0 A^0 A^1$), the listener can easily map a further intervallic expansion in A^2 and simultaneously deduce a mapping rule of augmentation to multiply by two (eighths augmented into quarters, m. 3). Because the first three

Figure 15 consists of two musical staves, (a) and (b). Staff (a) shows a sequence of motifs labeled A⁰, A¹, A^{2,a}, and B. A bracket under A⁰ and A¹ is labeled 'Diminution (plus sequence)'. A bracket under A¹ and A^{2,a} is labeled 'Sequence'. Staff (b) shows a sequence of motifs labeled A⁰, A⁰, A¹, A², and B. A bracket under the first A⁰ and A¹ is labeled '(repetition ... expansion ... augmentation and expansion)'. Below staff (b) are dynamic markings: *p*, *espress.*, *poco cresc.*, *legato*, and *dimin.*

Fig. 15. Diminution or augmentation plus full or partial sequence. (a) Bach, Fugue 15, *Well-Tempered Clavier*, 1, mm. 1–4; (b) Brahms, Piano Quintet op. 34, I, mm. 51–54 (*Allegro non troppo*).

motivic forms occur at a faster pace than the succeeding, slower paced one, listeners may hear A²'s similarity to its predecessors more strongly, as an asymmetry of longer back to shorter. Processing itself affects the perception of similarity, as Medin et al. (1993, p. 258) point out, and thus we may perceive similarity much more strongly in augmentation than in diminution. In any case, two rules are at work in Figure 15b as well—partial sequence (rule 5) and augmentation (rule 7), one of which embeds in the other:

$$A^0 + T_n(A^0 [d(v) \times q = d(v^1)]) = A^1 \quad (9)$$

(more simply: (S(d(A⁰)) plus expansion = A¹, where S is the transposed sequence of A⁰ and d(A⁰) the augmentation).²¹

Because augmentation reduces the sense of pace, whereas diminution increases it, one wants to know whether the effects of the formal asymmetry are directional—whether listeners detect a measurable decrease in similarity with diminution or an increase in similarity with augmentation (more about pace later).

Additivity in Both Melody and Duration

When listeners perceive augmentation or diminution (as in Figures 14 and 15), rule-mapped multiplication or division preserves the durational proportions. But other kinds of rule-governed stylistic inferences exist within a fixed metric framework, and these involve additivity in both melody and

21. It makes no difference which nests inside the other as the rules are commutative.

duration. In certain cases, adding a melodic tone often involves adding the same durational value. Thus additivity by rule in the two parameters becomes perfectly correlated. In short, if tempo and pace remain constant, the correlation allows the listener to construct only one cognitive rule. With respect to abstract generalization, such a cognitive rule is therefore very simple:

$$A^0 + c + w \rightarrow A^1 \quad (10)$$

(where c represents a continuation of A^0 's melodic implication; and given d = duration of a note and v = its value, w is such that $d(w) = d(v)$ for any $v \in A^0$).

Specific musical instances are considerably more interesting than this simple additive rule suggests. Consider Figure 16a. Here the motivic time frame is constantly expanded to accommodate the additive iteration. The notated metric additions in the score seem completely regular (2/4, 3/4, 4/4, 5/4), but cognitively the seamless grouping of each glide is less systematic (3, 7, 8, and then 10 notes). By the end of the 4/4 measure, the listener

(after m. 2 each rising line increases in length, first by one and then two tones)

(7 tones) (8 tones) (10 tones)

pp (begins as a sequence) (begins as a repetition) (begins as a sequence)

$A^0 \dots A^1 \dots$ becomes A^{2+} A^{2+} becomes A^3 A^4

A^0 A^1 A^2 A^3

p (addition to pitch plus subtraction from duration, resulting in a systematically faster pace)

Motives systematically lengthen while gradually quickening the pace before the sudden change of tempo

(♩ = 68) A^0, a A^1, b A^2, c A^3, d A^4, e (♩ = 88) A^5, f

ff *p* *p* *pp* *p*

(model sequence sequence additivity!)

(three notes) (four notes) (five notes) (six notes) (seven notes) (eight notes)

Figure 16 consists of three musical examples illustrating additivity. (a) Ravel, Suite from Mother Goose, IV, mm. 1-4. The score shows a sequence of rising lines that increase in length by one and then two tones. The tempo is marked 'très modéré'. (b) Rimsky-Korsakov, Capriccio Espagnol, IV, clarinet cadenza. The score shows a sequence of rising lines that increase in length by one and then two tones. The tempo is marked 'modéré'. (c) Beethoven, Symphony no. 1, IV, mm. 1-8 (Adagio). The score shows a sequence of rising lines that increase in length by one and then two tones. The tempo is marked 'Adagio'.

Fig. 16. Additivity. (a) Ravel, Suite from *Mother Goose*, IV, mm. 1–4 (*très modéré*); (b) Rimsky-Korsakov, *Capriccio Espagnol*, IV, clarinet cadenza; (c) Beethoven, Symphony no. 1, IV, mm. 1–8 (*Adagio*).

clearly deduces the rules of Ravel's game, having figured out that each rising line will extend in both melodic tone and eighth note. Moreover, in the last repetition, two rules—additivity plus sequence—subtly compete for attention.

Specifically, in the D-E♭ of measure 3 (start of the second bracket) one senses motivic return and thus projects an exact repeat (the second A^{2+}). Realization of this implication does occur, but then the extension of the rising line in the 4/4 bar denies the exact mapping. Starting in measure 4 (the 5/4), the repeat of the rising line begun on E♭—a tone higher—leads the listener to expect a sequence, which does come to pass. But now the listener also projects more than a simple sequence because she cognitively learned in measure 3 to expect an additional (ninth) note. The ensuing tenth tone is thus again a surprise (but perhaps less effectively so than it appears because a new melody enters in the oboe late in measure 4, vying for attention and hence somewhat obscuring the rule-governed additivity). The application of two rules—additivity and sequence—increases the amount of processing required and thus the level of aesthetic information.

In the less subtle though equally effective cadenza of Figure 16b, the rules governing the start of each melodic and durational grouping are uncorrelated. The triadic melodic tones and the number of notes increase with each varied swoop up and down, but simultaneously, durational values diminish, thereby increasing the pace. The melodic tones come in two groups—6 and 8 notes, and then 12 and 14 (not counting the start of each beat). Durationally, the four beats divide into groups of sixteenths (triplets and then faster quadruplets) and then shorter sixty-fourths (sextuplets and then septuplets). The durational value of each successively divided group is thus maintained within the beat. That the largest decrement occurs at the start of the sixty-fourth-notes reflects the aesthetic formal strategy of following A^0A^1 with further differentiation, except that here, of course, the increase only creates dissimilarity—an A^2 . Concerning the numerical increase of tones and the decrease in durational values, the listener perceives that systematic rules govern not only each two-beat group but also the pattern as a whole. One projects the pitch and durational variation according to the stylistic similarity (the four As).

As for rule, the realization of implied melodic continuation (c) adds tones while dividing the start of each group of durations:

$$A^0 + c + d(v)/q = A^1 \quad (11)$$

Of course, clarinetists do not perform the cadenza mechanically, in exact arithmetic notation. The change in pace is at first gradual and then noticeably accelerated. Nonetheless, the passage should be played strictly enough to allow the listener to rule-project the continuation accurately, thereby making the change at A^2 an exciting moment.

The opening of the last movement of Beethoven's First Symphony (Figure 16c) is probably the most famous case of additivity. The passage begins for all the world as if the motivic strategy here is to write ornamented sequences ($A^{0,a}A^{1,b}A^{2,c}$). At some point the listener realizes, quite surprisingly, that rules of additivity are at also work. Once deduced, each motive—from three to eight notes in length—exactly confirms the listener's rule-mapping as regards melodic ornamentation, the addition of pitch, and the decreasing durational values within each grouping. The gradual change from dotted cumulative notes ($A^{0,a}A^{1,b}$) to some isochronous ones ($A^{2,c}A^{3,d}$), the speeding up that takes place with each addition of pitch, and the fluctuation in dynamic from crescendo to piano and then to sudden pianissimo—all these tease the listener and whet one's sense of suspense until the cat is let out of the bag and the tempo-surprising allegro gets underway (Tovey, 1965, p. 24). Multiple play of this sort, which depends on the listener inferring the additive and sequential cognitive rules of the game at hand, is, of course, typical of the high syntactic wit found in the music of the Enlightenment. Although Beethoven's notebooks show that he could not multiply numbers, he understood the relevance of rule application as well as anyone when it came to music cognition.

Subtractivity and Additivity in Melodic Interval

Figure 17 shows that melodic interval can be the variable on which a cognitive rule is brought to bear. At the beginning, the melodic intervals generally decrease in size (octave, minor sixth, diminished seventh [=M6], minor sixth, perfect fifth, perfect fourth, major second, minor second). Much the same can be said about the beginning of measure 5 (in m. 6 both increase and decrease are present). Once the intervallic contraction gets un-

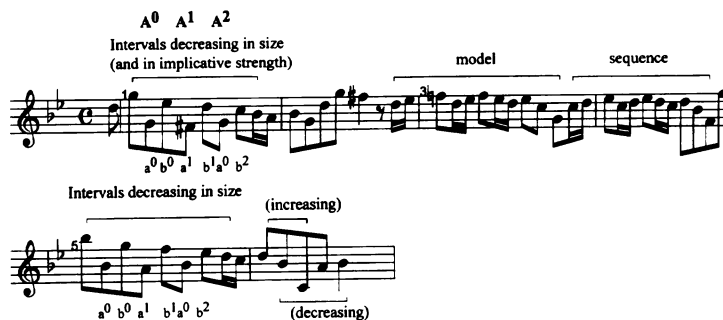


Fig. 17. Subtractivity. Handel, Concerto grosso in G minor op. 6 no. 6, IV, mm. 1–6 (Allegro).

derway, the varied stylistic repetition ($A^0 A^1 A^2$) becomes predictable. More specifically, because listeners hear individual intervals as implicative, they project a decrease in the degree of melodic implication.

Consider now intervallic additivity in the popular song of Figure 18a. The effectiveness of the high D results partly from cognitively deducing the systematic intervallic expansion (major second, minor third, major third, perfect fourth), which is to say, the gradual increase in the degree of melodic implication. The rule is not one that we can invoke on first hearing; awareness of it emerges retrospectively. Of course, in terms of extraopus style, a tonal pitch sequence of C-D-F-A usually goes to either a B \flat (if diatonic) or a C (if pentatonic)—thereby reaching the octave. Thus the surprising D contributes emotional affect because the general stylistic expectation is denied.

Figure 18b is a good deal more sophisticated. Apart from its superficial resemblance to the opening of Figure 18a (the unsuspected pentatonic ascent), the melodic rules operating here—three altogether—are complexly interwoven. (1) Each attempt at ascent ($A^0 A^1 A^2$) adds an extra tone; (2) the terminal interval of each motive expands until reaching the half-note G in measure 3 (just when it retrospectively dawns on the listener that more than one simultaneous rule is at work); simultaneously, (3) the pace grows quicker. The eighth-note patterns in the B-part (m. 4), whose falling implications suggest the melody will eventually descend, momentarily obscure the expectation of the high B \flat , but when realized, a pleasurable moment of release occurs. Both these Figures rely on the same kind of additivity. Even though stylistically differentiated, both share a cognitive rule of intervallic expansion, which ties them together.

Figure 18 consists of two musical examples, (a) and (b), illustrating intervallic expansion and additivity.

(a) Friml, *Rose-Marie*, "Indian Love Call," Refrain (Slowly). The notation shows a melodic line in G major with lyrics "When I'm call - ing you". Above the staff, a bracket labeled "Systematic intervallic increase" spans the first four notes (C, D, F, A). Below the staff, the notes are labeled A^0 , A^1 , A^2 , and B. The first note is marked *p*.

(b) Mahler, *Rückert-Lieder*, V, mm. 1–4 (Ausserst langsam und zurückhaltend). The notation shows a melodic line in G major with lyrics "(two tones)", "(three tones)", "(four tones)", and "!". Above the staff, a bracket labeled "Systematic addition of tones and systematic increase of pace" spans the first four notes. Below the staff, the notes are labeled (two tones), (three tones), (four tones), and !. The first note is marked *pp*. Below the staff, a bracket labeled "delayed realization" spans the last two notes.

Fig. 18. Multiple rules. (a) Friml, *Rose-Marie*, "Indian Love Call," Refrain (Slowly) [copyright held by Bambalina Music]; (b) Mahler, *Rückert-Lieder*, V, mm. 1–4 (Ausserst langsam und zurückhaltend).

What we learn, then, from Figures 17 and 18 is that listeners can abstract rules from parametric materials like melodic intervals. The bottom-up system, which processes melodic materials separately from learning (Narmour, 1990, 1992), attends to scaled properties of similarity and difference. So it does not care whether top-down, learned rules of organization are present. When, however, the top-down stylistic system encounters chromatic (C-C#-D-D# . . .), whole-tone (C-D-E-F# . . .), octatonic (C-D \flat -E \flat -E-F#-G-A-B \flat -C), or other systematic sequences of pitch, it will abstract cognitive rules that project the pitch continuation. In the cases of chromatic or whole-tone scales, cognitively constructing a rule is child's play: "Replicate the preceding interval." With the octatonic scale, the rule is more involved: "Alternate half- and whole-steps."

From this one might think that the major scale is also cognitively rule-governed, as two distinct tetrachords (whole-, whole-, half-step; whole-, whole-, half-step). But as many have pointed out, learned scale-step functions, which create goal-note stability on degrees 1, 3, and 5, can obscure this tetrachordal symmetry.²² Perhaps it is this "conflict" between bottom-up rule-mapping and top-down learning that has made the major system so long-lived: Implicit intervallic rule and learned pitch function vie for attention. In any case, projecting the continuation of the major scale probably relies more on simple mapping than on invoking tetrachordal expectation (the same seems true of the "natural," "melodic," and "harmonic" minor modes).

Cognitive Rules Without Stylistic Repetition: Other Parameters

Humans process information from both the bottom up and the top down. This is true for every level—from the primitive materials of individual parameters (expansions and contractions of pitches, intervals, durations) to style-structural repetitions (full, partial, or alternating sequences). Repeated durational patterns create specific event-structures via bottom-up processing while being holistically mapped as learned stylistic rhythms from the top down. Simultaneously, melodic and harmonic patterning may be formally differentiated and ongoing, devoid of any repetition. Music in fact teems with noncongruence between bottom-up and top-down processing (which is why music theory and music cognition are such difficult fields). The constant cognitive negotiation between the two aspects is what gives us an accurate and utilitarian sense of reality in the face of a dynamically changing context.

22. In tonal music, the asymmetry of the materials articulating the octave diverges further according to parameter: in harmony, steps I-II-V and I-IV-V occur more frequently as articulating nodes than I-III-V.

Although parametric materials constitute the primitives processed from the bottom up—and thus manifestly occur on lower hierarchical levels—bottom-up processing is recursively applied as higher hierarchical levels emerge. Thus bottom-up processing is not just relegated to lower levels. Likewise, top-down processing is not necessarily synonymous with higher-level events. Style exerts influence everywhere, from the basic elements that define parameters (like scale steps) to the hierarchical complexes that produce very high-level style structures (like common key relations). Because we have already discussed in some detail how cognitive rule and style-structural repetition operate together, let us now consider how iterative but largely unconscious rule-mapping can cognitively apply to other parametric style shapes.

TEMPO

Accelerando and ritardando are so utterly common in musical style, and thus so effortlessly mapped by the unconscious mind, that they seem to have been present from the very foundation of our musical memory. A listener tracks accelerando and ritardando by estimating the similar intervallic distances between successive durations through either a relative or constant shift in one's internal clock (Vos, van Assen, & Franek, 1997). Such bottom-up processing follows the Gestalt laws of similarity ($a^0 + a^1 \rightarrow a^2 \dots$).

One can also make sense of successive increment or decrement in tempo by top-down rule (many discussions about formalized cognitions of tempo change exist).²³ If, for instance, the rate of the tactus (foot tap) suddenly increases, the listener can deduce the multiple governing the change. Musical terms like *doppio movimento* (twice as fast) and *stretta* (immediately faster, as in an opera finale) are cases in point. Although solo performers rarely conceive of tempo change as mathematical behavior, those who conduct frequently describe it as calculus-like. Certainly the wide adoption of metronome markings for tempo change from the nineteenth-century onward indicates that composers thought of tempo in terms of numerical proscription.

In homophonic music, *rubato*, which means “robbing” in Italian, was frequently measured against a more-or-less constant meter. Many writers from the eighteenth- and nineteenth-centuries emphasize that when slight ritards follow slight accelerandos, the “borrowing” of time in the slowing down must be “paid back” by speeding up (see the discussion in Hudson, 1994).²⁴ Wagner (1897) makes this point in his writings on conducting,

23. See, e.g., Bowen, 1996; Clarke, 1989; Epstein, 1985, 1995; Feldman, Epstein, & Richards, 1992; Palmer, 1989; Repp, 1992, 1994, 1997a, 1997b; Sundberg & Verrillo, 1980; Todd, 1985, 1995.

24. For more historical discussion of tempo and *rubato* in the common-practice period, see Schuller (1997).

and the pianist-pedagogue Matthay (1913) illustrates the compensating principle of phrase rubato as a sinusoidal graph. Other writers, however, claim that what rubato borrows can be permanently lost. According to Hudson (1994, p. 321), McEwen's 1928 study measuring the note durations of piano rolls confirms this. The point is, the listener can make use of rules to project the degree of rubato, regardless of whether "compensation" takes place.

PACE

Listeners may also perceive gradual changes of pace as rule governed. Pace is independent of tempo and can affect either a single parameter or a group of parameters. For instance, in the melody of Figure 19, the accelerated pace from measure 2 to measure 3 moves on the manifest level in a clear ratio (from three notes per beat to four). Any listener tracking the tactus will perceive this.²⁵ Here one rule—sequence allied with crescendo—is thus replaced by another: change of pace (leading to the high C in m. 3, which realizes the triadic implications of the melodic pattern).

In Figure 19, the prolonged pedal point on the dominant hardly moves. In contrast, the melodic repetition in measure 1 divides the bar into two halves. Measure 2's sequence is also at the half-note level (but with more tones per half note and thus an increased melodic pace). In measure 3, the soaring melodic gesture to the forte on the dotted quarter note almost spans the whole-note level, despite the accelerated sixteenth notes. Thus, although pace here is parametrically differentiated, the listener's overall sense is distinctly one of carefully measured, and therefore rule-governed, acceleration.²⁶

25. One must distinguish the parameter of tempo (the rate of beat succession) from the parameter of pace (the sense of motion or activity relative to an established metric unit of time). The independence of pace from tempo is why it constitutes a parameter. Tempo may increase or decrease (in *accelerando* or *ritard*) because the listener's perception of tactus speeds up or slows down. But in both cases the amount of activity per beat may stay the same. Pace itself may also independently increase or decrease while tempo remains the same, accelerates, or ritards. Further, regardless of the durational level on which we perceive the beat, we may sense that pace increases, decreases, or remains unchanged. Both music theorists and cognitive psychologists are often careless about these terms, referring to pace when they mean tempo, and vice versa (see the clarifications by Cooper and Meyer [1960, p. 3] and Berry [1966, p. 317]). The possible relationships between pace and tempo are: same pace/slower tempo, same pace/faster tempo; slower pace/same tempo, faster pace/same tempo; slower pace/faster tempo, slower pace/slower tempo; faster pace/faster tempo, faster pace/slower tempo.

26. The term "harmonic rhythm"—where chord change speeds up or slows down—is really an aspect of pace. Hemiola—several duple groupings superimposed over triple meter—also involves an accelerated sense of pace that conflicts with the established triple meter at the bar level by a reduction of one third (e.g., $1/4$ subtracted from $3/4$ produces $2/4$). Hemiola thus also illustrates a rule-governed procedure.

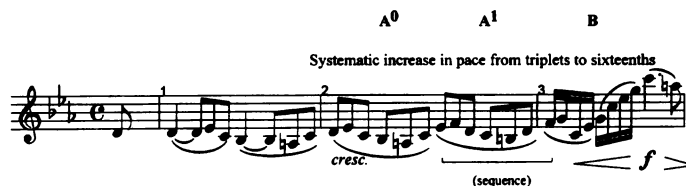


Fig. 19. Systematic increase in pace. Brahms, Sonata for clarinet op. 120 no. 2, I, mm. 36–38 (Allegro amabile).

Observe that the rule application here takes place in the differentiated part (B), unlike many earlier Figures where rule-making and stylistic repetition ($A + A . . .$) went together. To repeat: rule invocation is not dependent on similarity nor reducible to it (recall also in this regard the discussion of augmentation and diminution).

TEXTURE

Knowledgeable listeners can comprehend systematic additions to or subtractions from the voices of a musical texture. Rule-derived texture in variation form, for example, frequently increases in density and complexity by the systematic addition of tones (and thus durations) at the start of each new part. A first variation will occur in quarter notes, a second in eighths, a third in triplets, a fourth in sixteenths, and so forth; such melodic “divisions” are quite predictable and thus rule projective. In a fugal exposition (dux/comes), the stylistically predictable voice entries are texturally additive, as is the increased complexity that occurs during a stretto concluding a fugue. Figures 20a, c, and d display three incontrovertible cases of textural accretion perceivable by rule. Subtraction of textural layers can also be rule governed—for example, the waning of the last movement of Mahler’s Fourth Symphony and other postludal gestures where the density of voices systematically evaporates.

Of course, compositional procedures shaping one’s expectations about the gradual thickening or thinning of texture do not necessarily define where such changes will occur. Nevertheless, deduction influences the listener’s expectation concerning the gradual buildup or winnowing of textural complexity.

HARMONY

Increasing textural complexity usually results in increased dissonance. Thus, when voice entries increase, listeners may use cognitive rules to project a systematic pattern of accretion in the harmonic dissonance. In Figure

Harmonic fifths systematically added (along with increase in pace) ³

a

ppp

pp

increasing dissonance ... and decreasing

b

d5 d5 m3 m2 d5 m3 m3 d5
m7 m6 m2 m7 m7
P5 P5 P5 P5 P5

Exact retrograde, decreasing dissonance
and tempo, begins after the peak

c

For all our days are passed a way in thy wrath

For all our days are passed a way in thy wrath

For all our days are passed a way in thy wrath

For all our days are passed a way in thy wrath

Systematically increasing dissonance while accelerating tempo

d

f

traditional sonorities (sevenths)

consonances:	1	3	5	6	8	6	6	10	5	6	4	4	5	5	3	1
dissonances:			1	4	6	4	4	4	6	5	4	2	2	1	1	2

Dissonance: increase . . . decrease . . . increase decrease . . .

Fig. 20. Systematic increase in textural complexity and dissonance. (a) Ravel, *Daphnis et Chloé*, I, mm. 1–7 (Lent); (b) Britten, *War Requiem* op. 66, the harmonic progression ending the Requiem Aeternam, rehearsal letter 16 (molto lento); (c) Ives, Psalm 90, m. 60 (Adagio) [copyright held by Merion Music]; (d) Perle, *Songs of Praise and Lamentation*, III after Perle, *Twelve-Tone Tonality* (1977), p. 74 [copyright held by Boelke-Bomart].

20a, the additive stacking of harmonic fifths gradually and precisely increases the amount of dissonance present, and listeners will rule-map their expectations to anticipate the ongoing change (as well as to project the systematic growth in textural density).

The subtle variability in Figure 20b is similarly predictable. Calculated from all three intervals present in each chord pair (and ignoring octave doublings), the analysis shows that the harmony moves from open tritone (d5) to mild dissonance (m3, m7, and P5) to moderate dissonance (m2 added to m6 and P5) to strong dissonance (m2 and d5 added to P5; note the intensifying melodic leap to A, congruent with the incidence of maximum dissonance), then back to milder dissonance (m3/m7/ P5), and finally returning to the empty tritone (d5) that began the passage. Thus in terms of harmonic dissonance, Figure 20b has a kind of additive/subtractive quality about it, like an *accelerando/ritardando* or a *crescendo/diminuendo*. Unlike Figure 20a, however, here it seems likely that only on a second hearing would listeners prospectively apply a quasi-symmetrical rule projecting the dissonance arch. Both figures are through-composed (i.e., lack formal repetition).

Figure 20c is a more extended, famous case where “each note value progressively loses the length of a sixteenth note until the word ‘thy’” (Winters, 1986, p. 32). This subtractivity causes a systematically generated *accelerando*. Moreover, listeners project the accretion of dissonance here according to rule (at least up to a density point where the clusters become noise). They will also expect increasing textural complexity in the number of voices. In addition, the outside melodies produce rule-governed whole-toned inversions of each other until the last sonority, the apex of maximum dissonance. Indeed, the bottom two voices (low tenor, low bass) mirror the top two (high soprano, high alto). Rule-mapping the inversions would, of course, occur only retrospectively and probably after some learning.

In terms of dissonance, textural change, and registral expansion, this passage thus creates a symmetrical complex that is highly rule governed. The rhythmically congruent, contrasting outside voices together with the congruent dynamic change enable the cognitive prediction of an ever-expanding “wedge” (Winters’s term). All the parametric changes typify $a^0 + a^1 \dots$ and create ongoing material similarity and predictable variation but without any formal repetition. Once the melodic peaks of the outer framework occur, an exact subtractivity begins in the same parametric way (after the vertical arrow, not shown). With some exposure listeners will learn to predict the palindrome, which is to say, project a “wedge” that operates both vertically and horizontally (like the “hairpins” of a *crescendo-diminuendo*). In sum, Ives uses five rules to generate the passage, and, listeners can project most of these, though certainly not all on first hearing.

Such rule deduction is one reason why rehearing is so enjoyable: We convert our retrospective evaluations into prospective ones and thus unconsciously retest the rules that we initially learned (while also perceiving new ones that we missed on earlier occasions). In other words, through learning, rehearing music changes retrodiction to prediction. In philosophical and aesthetic terms, rehearing music combines the rational (rule deduction) and the empirical (rule testing) into the beautiful.

Figure 20d is another instance of additive dissonance, but one where precise cognitive rules are not so easy to formulate. As in Figure 20b, I have vertically calculated the harmonic intervals between every voice in each sonority (omitting octave doublings, of course). Because of the harmonic complexity, the integers here refer to numbers of dissonances and consonances rather than to the intervals themselves (as in Figure 20b). Beginning with a consonance, the passage grows increasingly dissonant but then becomes somewhat less so toward the end, yet it is not clear that listeners would discern an underlying rule to cover the change. Still, one does perceive a general arch in the dissonance and deduces that some governing principle is at work, however unsystematic.²⁷ As in the Britten and the Ives examples, Perle's passage also grows texturally more and then less complex. Again no stylistic repetition (AA . . .) accompanies the rising and falling curve of the dissonance.

It is not certain in the four cases of Figure 20 how common this treatment of dissonance is. Indeed, in the extraopus sense dissonance in tonal style tends to remain a localized phenomenon. The reason is that dissonance per se is not easily wrested away from bottom-up processing. Thus it does not lend itself well to abstraction, a necessary condition for rule-mapping. Consequently, listeners are not particularly good at rule-mapping additivity or subtractivity in dissonance, particularly in the face of complex atonal sonorities.

Even tonal harmony resists such abstraction. Tonal passages involving extensive dissonance always involve other variables—chord function (e.g., circle of fifths), voice leading (parallel vs. contrary motion), scale step, tonicization, modulation, chromaticism, and so forth. These top-down, learned aspects determine the overall, complex sense of tonal harmonic implication, including the function of tonal dissonance. Soprano position, bass position (inversion), and common-toneness between chords—putatively bottom-up properties—also affect the perceived quality of dissonance. Listeners are able to project harmonic continuation accurately in tonal music, but that syntactic ability relies heavily on learned mapping, not on iterative rule-mapping of the types discussed earlier. In tonal harmony, ongoing in-

27. Perle's book does not discuss the aspect of dissonance in this example; the analytical observations regarding their cognitive reality are mine.

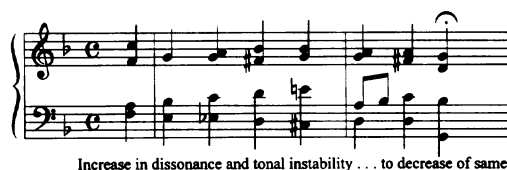


Fig. 21. Harmonic syntax contrasted with rule-mapped dissonance. Bach, “O Ewigkeit, du Donnerwort” (Chorale), fourth phrase (text omitted).

crease or decrease in dissonance is only one among many variables that influence the listener’s sense of harmonic expectation.

Consider Figure 21. Because scale step plays a crucial role in the syntax, it is not just the presentation of dissonances in Figure 21 that enables the listener to project the progression. Beginning with a tonic chord (in F) which is completely consonant, the harmony moves to a diminished chord (one dissonance, a tritone), a second-inversion seventh-chord (two dissonances, a tritone and a second), an augmented chord (no dissonance but unstable nonetheless), a diminished seventh (two tritones), a dominant chord in G minor (with one dissonance, a major second/minor seventh that suspends a fourth over the bass), a dominant-seventh chord (two dissonances, a minor seventh and a tritone), and finally a tonic in G minor (no dissonances). Thus the voice leading (descending linear bass, contrary ascent in the tenor until the dominant, and a curved soprano melody), the modulation (supported by unforeseen chromaticism), the changing scale-step functions, and the tonicizing of G minor contribute strongly to the sense of harmonic instability or stability as the sequence of dissonances. Indeed, the tonal factors are probably more salient than any rule deduction concerning dissonance. The listener familiar with Bach chorale style clearly recognizes the direction of the harmonic process, but that expectation results mostly from the learned mapping of tonal syntax, which in this style relegates dissonance to the foreground rather than to abstract generalization.

Mirror Forms: Inversion and Retrograde as Rule-Mapped Symmetries

In earlier discussions, we distinguished simple stylistic mapping from cognitive rule making—from the iterated increment and decrement that often accompany formal similarity. Although we have concentrated mostly on projection, prediction, and prospection, the possibility exists that listeners may retrospectively deduce cognitive rules to arrive at contextual meaning.

Melodic inversion is a case in point. Mirror form generates a nontransitive, reflexive perception that depends on symmetry between varied copy and model. When the listener first encounters inversion, the cognitive projection of the rule will be entirely retrospective, after the fact. Its mapping will therefore be retrodictive, not predictive. Retrospection characterizes the tense of the cognitive rule, so to speak: “What *was* just heard is based on a model whose transform replaces ascent with descent or descent with ascent while preserving intervallic sameness or similarity.” Of course, if listeners hear a given inversion a sufficient number of times, then they will learn to expect it and map the continuation simply as a repetition.

Because melodic inversion operates only on ascending and descending direction (lateral motion is unaffected), the operation can take place by just changing the sign, as it were. Thus the “plus” (+) of ascent becomes a “minus” (–) or down, and the “minus” of descent becomes a “plus” or up. If an inversion is exact, the change in direction has no bearing on the melodic intervals, the size of which will remain the same (but of course in the opposite registral direction). In tonal melody, however, composers usually allow the size of the interval to be determined by the diatonic set (so perfect fifths become perfect fourths, major intervals become minor ones. etc.).

In simple inversion (without transposition) the rule takes the following symbolic form:

$$-\text{reg}(A^0) = I \quad (12)$$

Or:

$$-\text{reg}(A^0) = A^1 \quad (13)$$

where $-\text{reg}A^0$ (a change in sign) denotes the change of each registral motion to the opposite direction while the intervals of A^0 remain constant (I stands for “inversion”). Note here the equals sign (=) instead of the arrow (\rightarrow), which symbolizes that, at least on the initial hearing, the output is a retrospective assessment rather than a prospective one.

When both transposition (T) and inversion (I) occur, two rules operate simultaneously:

$$A^0 + T_n[-\text{reg}(A^0)] = TI = A^1 \quad (14)$$

One can also express this with registral change nesting inside the transposition:

$$T_n[-\text{reg}(A^0)] = A^1 \quad (15)$$

Because invoking two rules increases the amount of cognitive processing, short passages heard as inversions will often be construed as simple retrogrades because simple retrogrades require only backward mapping, not rule. For example, in a short up/down pattern like C-D-E-D-C, one

could hear E-D-C as a transposed inversion of C-D-E up a third (T_4I). But it is easier to hear E-D-C as a simple retrograde.²⁸ The limitations of short-term memory, however, which must store the whole pattern before temporally running it backward item-by-item (as opposed to spatially processing an array), constrain our cognitive abilities to retrograde long patterns. But short-term memory constrains inversion as well. Hence syntactic rules are efficient in projecting inversion and retrograde only when the transform's length is manageable. That said, short-term memory in general deals with inversion better than retrograde (Dowling, 1972) because in inversion at least the temporal order of events remains analogous. Yet because mirror forms tend to be difficult to apprehend when extended, they probably have categorical properties that have to be satisfied in order to be perceived.

Deducing a discontinuous inversion requires an even higher degree of functional correspondence. If too great a distance exists (Dowling & Bartlett, 1981), if the parametric variability of the intended inversion is too great (Francès, 1957/1988, Krumhansl, Sandell, & Sargeant, 1987), or if the context of the inversion is incommensurate with the model, then listeners are unlikely to connect transform to model and thus construe an inversion.

Consider the contextual effects of tonal function. If simple pitch relations and intervallic size are held constant during a change of register, contour by itself will create a pattern that one may cognitively deduce as a mirror. In C major, a pattern like G-A-B-A-G (up/down) sounds symmetrical because the midpoint B, as the leading tone, both ends the ascent and begins the descent, thus establishing an axis. That in this case scale step is functionally congruent with the reversal of contour and that all intervals are major seconds easily allow one retrospectively to perceive a symmetrical retrograde.

But when intervallic variation enters the melodic mix and alters scale-step sequence, the symmetry is no longer quite so salient. In a tonal pattern like C-D-F-E-C (up/down), F may appear to be the perceived axis, but because both intervallic sequence (M2-m3-m2-M3) and scale-step sequence (tonic, supertonic, subdominant, mediant, tonic) vary slightly, the cognitive sense of symmetry is somewhat weakened. This is true not only because C and E are goal notes (tonic, mediant) and D and F are nongoal

28. The degree to which retrograde emerges in simple patterns like the one above depends on the amount of transposition. If the pitch that starts the transform is the same as, or close in register to, the terminal tone of the original (as above), then hearing a retrograde is more likely than hearing an inversion. If the pitch is closer to the initial tone of the original, then perception of inversion will occur (e.g., an ascending pattern of C-E-G followed by a descending D-B \flat -G, where D is only a step away from C). Thus in simple patterns the locus of the axis of symmetry determines the rule applied. That pitch distance between the initiation of the transform can vary vis-à-vis either the starting or ending tone of the original is partly why theorists sometimes mistakenly call patterns that are retrogrades inversion, and vice versa.

notes (super tonic, subdominant) but because the D-F skip may sound like a gap with the subsequent E filling it in. The net effect is that the ascending minor-third skip tends to articulate the descending E as a structural tone, hence moving the axis away from F and thus obscuring the sense of retrograde (or transposed inversion). Of course, if the E were flatted, perfect intervallic symmetry would occur (M2, m3, M2, m3), but even in C minor the scale-step difference between D-F and E♭-C attenuates the perception of a mirror (whether retrograde or inversion). In a word, perceiving musical symmetry is highly dependent on context.²⁹

Inversion may operate on a pattern of any size in the musical fabric, from single interval to extended passages. Figure 22a–b illustrates some typical symmetrical inversions that are easily perceived. Although in Figure 22a the cello introduces the motivic model, which is immediately followed by the viola's transposed inversion (nearly exact; the cello's surface chromaticism is not mirrored), the first phrase ends on the pitch (C) that begins the symmetrical second phrase. So two rules are at work: inversion and transposition (up a second). The dovetailed phrasing establishes a prominent axis of symmetry despite the change in harmony (note that rhythm prevents our perceiving the passage as a retrograde).³⁰

In Figure 22b, the inversion mirrors the first phrase almost exactly; again transposition is involved (down a minor third). The harmony changes in the second phrase, small deviations occur in interval and dynamic, and the beginning of the inversion shares no axis because the closed half-note in measure 4 separates the two phrases. Nevertheless, the exact reflection of contour in the inversion and the exact durational patterning ensure the perception of the mirrored symmetry.³¹

29. In a general sense, retrograde and retrograde inversion are difficult to hear and thus to backward-map because all initial functions become terminal and all terminal functions initial. This puts an enormous load on memory as regards rule-mapping. For these reasons, the perception of extended retrograde is all but impossible. In fact, retrograde is very rare except in short motives (Berry, 1966, p. 397). In a lengthier context, retrograde tends to be purely a constructive device, affecting the musical content but only rarely being heard. Composition obviously relies on rule applications, but such rules are not necessarily equivalent to cognitive rule uses, Ives's palindrome earlier (Figure 20c) notwithstanding. I shall return to retrograde, constructivism, and cognitive rules in the next section.

30. For partial inversions that form parts of antecedent-consequent phrases, see Haydn, Symphony no. 100, III (trio) or Symphony no. 104, I (principal theme). The variation of the inverted parts (consequents) underscores the argument made earlier in connection with Figure 13 about formal differentiation and closure. For a nineteenth-century example, see Mendelssohn, Symphony no. 5, II.

31. Using this very melody (Figure 22b) as a test item, Rosner and Meyer's experiment (1986) produces data showing that listeners do not perceive its "underlying process" as a "complementary type." The authors express puzzlement by this, suggesting that the melody may be a "devious stimulus." It is possible, however, that the task confused the participants in that they were asked to rate the examples concerning whether they were "good or poor examples of the underlying process," not whether the melody was an inversion per se.

Figure 22 consists of three musical examples labeled (a), (b), and (c).
 (a) Schumann, Piano Quintet, I, mm. 57-60. This example shows two staves. The top staff is for the viola, marked *mf*, and the bottom staff is for the cello, marked *mf espress.*. A bracket labeled "inversion (and transposition)" spans the two staves, indicating a relationship between the two parts.
 (b) Brahms, Symphony no. 1, III, mm. 1-10. This example shows a single staff with a *p* dynamic marking. A bracket labeled "inversion (and transposition)" is placed below the staff.
 (c) Brahms, *Ein deutsches Requiem*, "Wie lieblich," mm. 1-8. This example shows two staves. The top staff is labeled "Introduction:" and has a *p dolce* dynamic marking. The bottom staff is labeled "Wie lieblich (full chorus)" and has a *p dolce* dynamic marking. A bracket labeled "inversion (and transposition)" spans the two staves. Below the bottom staff, the lyrics "sind dei-ne Wohn-ung-en," are written.

Fig. 22. Inversion and transposition. (a) Schumann, Piano Quintet, I, mm. 57–60 (*Allegro brillante*); (b) Brahms, Symphony no. 1, III, mm. 1–10 (*Un poco allegretto e grazioso*); (c) Brahms, *Ein deutsches Requiem*, “Wie lieblich,” mm. 1–8 (*Con moto moderato*).

This cannot be said, perhaps, about the inversion in Figure 22c, because here timbral, harmonic, and functional differences between the orchestral introduction and the thematic inversion sung by full chorus are much greater. These make the symmetry harder for the listener to construe and thus make retrospective rule invocation less likely.³² Moreover, the transposition is large—down an octave and a fourth. By contrast, the melodic symmetries in Figure 22a–b were more salient so that, once deduced, listeners could readily map both rules of inversion and transposition.

Visual Inversions Versus Musical Ones

Both inversion and retrograde retrospectively generate isomorphic structures. Only because listeners can reconstruct the original model from the transformation (Hofstadter, 1979) can they construct a mirrored symmetry between the parts. For this reason, inversion emerges most clearly when

32. I came to this conclusion because, although I have listened to (and conducted) this piece on numerous occasions in the past 20 years, I never noticed the transposed inversion until I saw it mentioned in Crist, DeLone, Kliever, Rowell, & Thomson (1982, p. 91), this despite my strong interests in relating analysis to performance.

model and transform appear directly opposite the axis. That is, inversion is most salient when it contiguously follows its model and when a fairly high degree of parametric—and thus contextual—similarity is present.

In vision, mirrored constructions normally proceed automatically and unconsciously and seem to be built into our operating systems (Uttal, 1988, pp. 144–146). We can, for instance, recognize a familiar face no matter how it is rotated or inverted (Bruce, 1991, p. 24), and, indeed, we can reconstruct known visual forms by viewing only the left or right or top or bottom of the axis. As Deutsch (1999, p. 362) says about vision, “We must have evolved mechanisms that preserve the perceptual identities of objects regardless of their orientation relative to the observer.”³³

However, concerning the cognitive rule-mapping of inversion in a temporal, syntactic medium like music, our mental processing is not nearly so facile, even though twentieth-century composers have mistakenly assumed that visual transforms and various mirror operations in music are cognitively equivalent (see, e.g., Schoenberg, 1950/1975, who was also a gifted painter). Indeed, ongoing interactions among various parameters may obscure the musical identification of the corresponding constituents such that no axis of symmetry emerges. If in vision symmetrical parts are “identical in shape but opposite in spatial orientation” (Arnheim, *Visual Thinking*, p. 63), then the crux of the matter in music lies in how cognition mentally “identifies” and defines the “orientation” of the mirror.

To explore the analogies and differences between visual and musical mirrors, consider the synthetic illustrations in Figure 23. In the first case (Figure 23a), a simple representation divided into four triangles, the viewer has no trouble in seeing both the horizontal and the vertical axis. The second case (Figure 23b) presents an analogous musical example; although the size of the intervals slightly varies, owing to the tonal context, one can easily hear both a vertical axis on A and D and, at least by analogy, a horizontal one on F, the beginning and ending tonal center. In Figure 22c, one triangle mirrors another only at the vertical axis, but the visual symmetry between model and mirror is still readily apparent. However, in the musical analogue to this (Figure 23d) the symmetry is less clear; for instance, rather than hearing a vertical axis emerge on the second F, one could construe the descent from A to D as a simple linear pattern, overriding any articulation on the middle note (F). Still, it would not be unnatural to hear the second part of the melody as an inversion (mirror) of the first. The visual symmetry of Figure 23e has two vertical axes along with one horizontal one, again easily seen. Figure 23f is the contrapuntal analogue (again with two axes, which are stemmed). So far, then, the musical examples and the visual figures appear reasonably analogous.

33. Perceiving short inversions and retrogrades, however, may in fact be built into the auditory cognitive system. More research is needed to determine this.

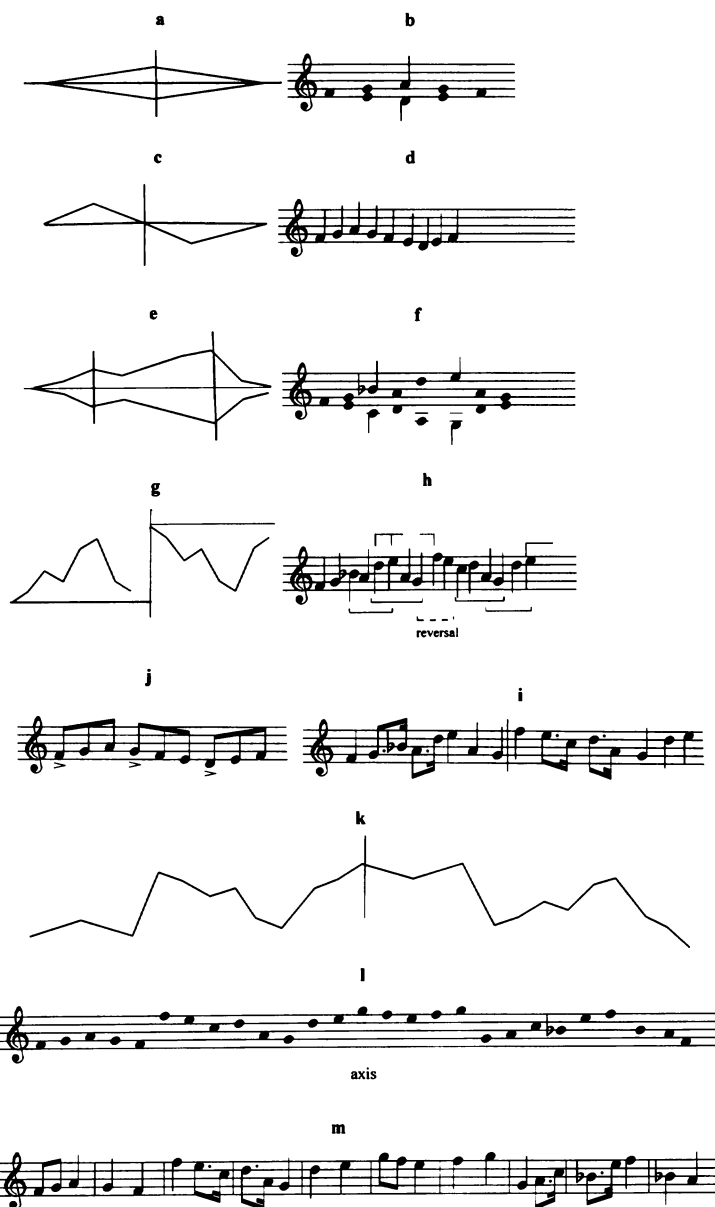


Fig. 23. Analogues of visual and musical inversions (see text).

Now consider Figure 23g. Here a visual form is mirrored after a vertical axis (a horizontal line on the bottom and the top of each segment enhances orientation). The spatial separation between model and copy—a large transposition, as it were—does not prevent our seeing the symmetry. Figure 23h is the melodic analog (bounded by the staff lines), with the inversion trans-

posed up an octave (the axis on the high F is unmarked). Although the symmetrical inversion in the visual form (Figure 23g) is quite apparent, the inversion in the musical analogue (Figure 23h) is quite difficult to hear. This is because time envelops syntactic processing (moving only left to right, as it were). Consequently, in Figure 23h the ascending leap of the seventh (G-F) at the axis, which implies a reversal (dashed bracket), assimilates the immediately following F-E-C, and this latter pattern conceals the start of the transposed inversion (and thus the axis). Moreover, the mappings of motivic similarity (regular brackets) obscure our hearing the beginning of an inversion on the high F. Further, the emergence of a discontinuous realization from the leading tone to the tonic (the rising D-E to the high F) blurs the mirror (arrowtail). Finally, the last note on the leading tone (E) leads the listener into projecting a continuation to the tonic F. Of course, one still has some sense that orderings of pitch and interval are consistent in some way.

But a vague sense about consistency of melodic content is quite different from retrospectively deducing a transformational rule and then mapping an inversion. To do this, one needs additional cognitive cues. For this reason, composers almost always rely on durational, metric, or harmonic patterning when they want to bring an extended inversion to the cognitive fore. And plenty of evidence demonstrates the importance of rhythm and meter in the perception of musical transformations (see Dowling, 1994; Halpern, Bartlett, & Dowling, 1998; Jones, 1987; Jones & Ralston, 1991). Figure 23i, which has the same pitch sequence as Figure 23h, shows how effective differentiated rhythms are in establishing a functional axis of symmetry (vertical line).

By the same token, parametric interference can prevent the perception of melodic symmetry. The dynamically stressed accents in Figure 23j, for example, create metric groupings that obscure the potentially mirrored pitch ordering found in Figure 23d (on accent structure and recognition, see Jones & Ralston, 1991). So parametric context can either facilitate or hinder the cognitive perception of melodic inversion.

Figure 23k's mirror configuration is more difficult to see than any of the previous visual examples. Yet with the aid of the vertical line, we can unravel the inversion rather quickly. Figure 23l shows the musical analogue, but sensing the inversion here is a lot more difficult. Indeed, it has all but disappeared. The distance of the transposition (M9), the lack of other parametric cues, the sheer length of the sequence (26 tones), and the absence of closure and thus a clear sense of form completely prevent the emergence of an axis. In Figure 23m, rhythm and meter help delineate the symmetrical relationship, but even here the heard connections between the dotted rhythms are more salient than those between the patterns of eighth notes and quarter notes (mm. 1–2, 6–7), which actually divide the form in half.

Only with a good deal of repeated exposure can the experienced listener begin to sense that this melody has symmetrical properties.

What we conclude from Figure 23 is that, in general, musical mirroring tends to be more difficult to perceive than its visual counterparts. Listeners can and do deduce rules to recognize melodic inversions but only if the patterns are short. Longer ones require a context suffused with other parametric cues. More experimental testing along the parametric lines suggested here remains to be done if we are to determine fully the perceptual conditions of melodic symmetry (Krumhansl et al., 1987). We simply do not know the extent to which rule-mapping of inversion takes place in complex, extended contexts.

Inversion and Affect

One can find inversion on all levels of melody, but, as we have seen, listeners most readily hear mirroring in rather small units. In Figure 24a, the descending interval ending the second phrase ($E\flat-B\flat$) exactly inverts the ascending fourth ending the first phrase ($E\flat-A\flat$). Because of the exact rhythmic conformance, the listener easily recognizes the discontinuous intervallic relationship and unconsciously deduces that transformation is at work. The invoked rule allows one to hear the syntactic playfulness between the two cadential gestures. Figure 24b is similar except that the inverted interval is tonally “adjusted” (P5 becomes P4); nevertheless, because melody and rhythm repeat exactly, the listener cannot avoid perceiving the symme-

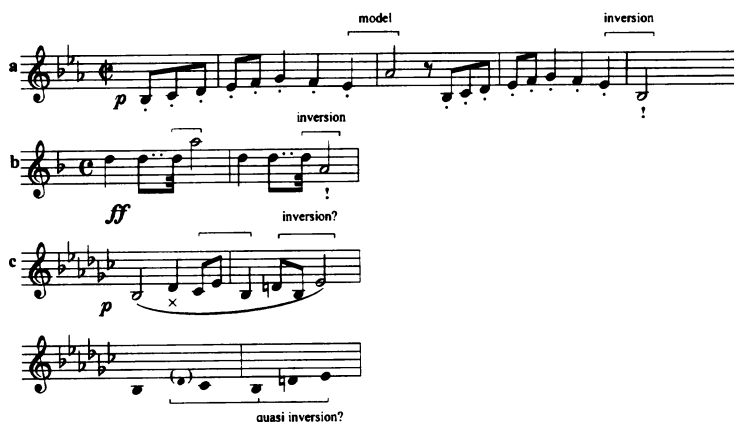


Fig. 24. Inversion and quasi-inversion. (a) Beethoven, Piano sonata op. 10 no. 1, III, mm. 17–20 (Prestissimo); (b) Haydn, Symphony no. 104, I, mm. 1–2 (Adagio); (c) Stravinsky, Suite from *The Firebird*, Berceuse, mm. 1–2 (Andante).

try (later in the introduction, when the inversion recurs, the second interval becomes a perfect fifth, to the surprise of the listener).

Although musical inversion brings about a contrast in contour, most writers believe that registral symmetry creates qualities of balance, proportion, assimilation, and mutual completion. Meyer (1973, p. 175), for instance, speaks of melodic symmetry in terms of complement, and Jones (1974) shows that people do perceive some patterns as complements. Consequently, a slight denial of symmetrical rule-mapping may give retrospective pleasure. Consider Figure 24c, whose motivic mirroring is quite concealed. The attentive listener senses (in retrospect) that some kind of rule-governed content is present with the eighth-eighth-half-note (m. 2), but because the higher-level structure underpinning the motives is not exact (see the brackets), the inversion is elusive. Solso's (1994, p. 146) remarks about symmetry are apt: "A *perfectly* balanced painting would be rather boring. . . . While we understand order, we find minor visual dislocations interesting and invest greater effort in investigating them." With reference to music cognition, the issue is the extent to which such interesting dislocations deform the symmetry and thus prevent the perception of inversion altogether. Figure 24c sits very finely on that edge.

Hierarchical embedding of melodic inversion may make rule-based mapping less likely, not necessarily because listeners are unable to compare different levels but because manifest events may obscure the similarity made covert by level differentiation. This was the case in Figure 24c and is also true of Figure 25a. Although low-level symmetry is manifestly absent, we can analyze the quarter-note level as an inversion—an ascending perfect

Fig. 25. Quasi-inversion. (a) Dvořák, Cello concerto, III, mm. 33–36 (Allegro moderato); (b) Mozart, Sonata K. 284, III, Variation 1, mm. 13–17.

fourth of F#-B followed by a descending perfect fourth of F#-C#. But is the descent heard as an inversion? One thinks not because of the lack of surface similarity between the two motives, whose low-level forms are AB. But if not symmetrical, then what in Meyer's term melodically accounts for the sense of motivic complement in this example?

Of course, surface difference need not obscure the cognitive recognition of inversion. The first full bar of Figure 25b displays a partial sequence on beats 2, 3, and 4, with intervallic expansions generally like those seen earlier (recall Figures 6–12). The second, third, and fourth beats of measure 2 then invert the pattern of measure 1, except the registral direction at the end of each motive stays the same—up ($A^{3,a}$ $A^{4,b}$ $A^{5,c}$). This “dislocation” subtly changes the cognitive rules of the game, but the playfulness does not obscure the listener's sense that inversion (or rather quasi-inversion) of the original intervallic expansion is relevant to the mapped expectation (note in m. 3 that the analogous beats are purely sequential).

Retrograde Inversion

As for extended inversion, in the history of music, composers have used it more than retrograde. Because contour has a strong effect on serial order recognition (Boltz & Jones, 1986; Boltz, Marshburn, Jones, & Johnson, 1985), training can make feasible the perception of longer inversions, particularly if other parameters such as rhythm and meter define them.

What of retrograde inversion? Short ones are probably more difficult to perceive than short inversions or short retrogrades (Dowling, 1972). The reason is that inversion retrograde requires multiple cognitive rules. Consider the motives in Bach's fugue subject in Figure 26. One might say that the second eighth-note motive of C-D-G (top staff; ascending M2, descending P5) retrogrades and inverts the preceding motive of C-G-A \flat (ascending P4, descending m2 following the two opening sixteenth notes). But does the listener retrospectively hear this mirrored symmetry? Although replacement of a descending fourth by a descending fifth is a common variant, the scale-step substitution of the dominant G (downbeat of m. 2) for the submediant A \flat (m. 1, beat 3) clearly obscures the connection. More important, on the quarter-note level, partial sequences of intervallic expansion (bottom staff) create the structural tones of a linear process on the half-note level (first staff, arrows). Thus any ascription of retrograde inversion here seems quite rationalistic in terms of the listener constructing a rule. Rather, it is the varied motivic repetition (A^b) that creates the subtle form of play.

The form, however, calls into question the implicative status of the higher-level descending line of A \flat -G-F (first staff). The first segment of each mo-

The figure shows a musical score for the first three measures of Bach's Fugue 2. The notation is in G major (one sharp) and 4/4 time. The first staff (treble clef) contains the first measure, with a bracket labeled A^0 above the first two notes (G4 and A4). Above the next two notes (B4 and A4), there is a bracket and the text " A^1 or A^b ? Retrograde-inversion or simple ornament?". The second staff (treble clef) contains the second measure, with three brackets labeled A^0 above the first three notes (G4, A4, B4), followed by three exclamation marks (! ! !). The third staff (treble clef) contains the third measure, with a similar rhythmic pattern. The fourth staff (treble clef) shows a sequence of notes labeled A^0 , A^1 , A^2 , and B, with brackets below them labeled "(partial sequence)" and "(partial sequence)", and the text "expanding intervals" below the entire staff.

Fig. 26. Quasi-retrograde-inversion. Bach, Fugue 2, *Well-Tempered Clavier*, 1, mm.1–3.

tive begins rhythmically and melodically the same—two sixteenths to an eighth (second staff, $A^0A^0A^0$). This would seem to send a message to the listener that stylistic mapping is to be straightforward. But with the third A^0 (after the downbeat of measure 2) the listener does not stylistically know what continuation to map. Is the tone to follow the third C-B-C supposed to mimic the first motive (A^0) and leap down? Or is it to mimic the second motive (A^1 or A^b) and step up?

If the step up to D in the second motive creates a retrograde inversion, then the resultant $A^0 + A^1$ would imply an A^2 . If it creates ornamental variation (A^b), then that would support a less-specified projection—simply that change will continue. However, because in the face of ambiguity listeners tend to opt for a mapping based on contiguity rather than on discontinuity (Tenney with Polansky, 1980), we probably hear the third motive's move up again to D as another surprise, which exactly repeats what preceded it. That is, sameness occurs instead of a change. This formal twist thus maintains the suspense, yet that is immediately followed by still another surprise: a larger leap downward plus two sixteenths now falling on the beat (hence now $A^0A^1A^2$, third staff). The rhythmic change jerks us forward, and by an unforeseen, syncopated return to the quarter note A^b (m. 2), deflects and delays the completion of the linear descent to E^b . In sum, the emergence of the implied, descending higher-level A^b -G-F (first staff) is no simple matter because the formal expectation ($AA \rightarrow A$) is somewhat ambiguous. Hence the listener does not completely anticipate the realization of the skip down to F.

Subtractivity and Additivity in Musical Form: The Importance of Learning

Normal listeners may learn to deduce rules governing formal manipulations. Consider the phenomenon theorists have variously referred to as “foreshortening,” “telescoping,” “elimination,” or “successive segmenting,” in which formal units become progressively shorter. Figure 27, from the famous first movement of Beethoven’s Fifth Symphony, is a clear case. Throughout the passage, there is unrelenting eighth-note motion, yet each repeated subphrase is successively halved: following two four-bar phrases (the second a slightly varied repetition, $A^{0,b}$), the third and fourth phrases reduce to two-bar lengths (B^0B^1), and these in turn are succeeded by one-bar motives (C^s).

Listeners can absorb and project this formal “compression,” even though Beethoven’s music requires other kinds of attentional processing. As the formal units grow shorter during the crescendo (mm. 10–13), harmonic change increases apace and grinds against the dissonant pedal points and

Successive shortening of phrase lengths:

4 A^0 4 $A^{0,b}$

2 B^0 2 B^1 (1) C^0 1 C^1 1 C^2

1 C^3 1 C^4 1 C^5

(melodic sequences)

Fig. 27. Systematic formal division. Beethoven, Symphony no. 5, I, mm. 25–44 (Allegro con brio).

insistent sforzandi (mm. 14–19) until the climax in measure 20. The inexorable sense of all this instills a tremendous anxiety in the listener until the release of the explosive forte. Compression speeds up the amount of information to be processed, causing the listener to remain on high alert.

Expert professional listeners (musicologists, theorists, composers) and other highly committed music lovers are cognitively able to step back from the intense processing of Figure 27 and (1) identify, remember, and compare the formal lengths involved, (2) analyze the precision of the formal construction, and then (3) formulate a predictive rule (“expect each motivic change to decrease in length by one half”). We have already mentioned the phenomenon in two other instances (recall the successive halving of durational values that created shorter motives in Figure 5b and the similar one-bar segmenting of mm. 17 and 18 in Figure 12b, following the two two-bar motives of mm. 13–14 and 15–16).

The converse of Figure 27, as in the cumulative motives of Figure 1 earlier (by Beethoven), also follows a systematic formal pattern: One-beat motives occurring four times are succeeded by two-beat motives occurring twice, which in turn are succeeded by the final motive that takes up an entire bar (four beats). In Classic music, such formal cumulation is utterly common. Another instance (by Mozart) occurs in Figure 13 where both the antecedent [AAB] and the consequent [AAC] phrases each create a phrasing of $1 + 1 + 2$. Shortening and then lengthening the form in multiple units is also common. In Haydn’s melody of Figure 11a, the first two motives last two beats, the next three are one beat each, and the last two each span one bar.

To repeat: Ordinary listeners can learn to perceive successive forms of either cumulative multiples or countercumulative divisions as rule-generated—as any college music teacher will attest.

Compositional Rules Rather Than Cognitive Ones: Proscriptions Generated from the Top-Down

“Learned style” is a well-known phenomenon among music scholars because serious composers have always shown interest in intellectual craft and, in the hopes of creating unity, coherence, consistency, integration, and the like, have frequently relied on rule proscriptions to generate sophisticated, complex music. I do not refer here to the workaday compositional rules of tonality so well codified they have become a mainstay of the undergraduate curriculum—for example, the harmonic rules of part writing (voice leading, voice doubling, etc.), the rules of form (governing sonata, minuet, rondo), the rules of orchestration (e.g., applying effective instrument

doublings), and so forth. Rather, I mean “pure,” top-down construction that comes to light only with objective analysis, conscious thought, and repeated study—the kinds of “hidden” compositional rules that scholars uncover and teach in concentrated analysis classes. Such constructivism relies not on perceivers to deduce and project a general cognitive rule but perforce to think consciously about a specific compositional one generatively imposed on the music.

Indeed, the aesthetic of some types of contemporary music requires listeners to consciously cogitate while they process the input. Listening habits, however, can be quite uncongenial to such pure rules. Most music lovers are generally unprepared to think about covert constructive rules because the normal habits acquired in listening to the common repertory (tonal music) do not entail apperceptions of this type. Consequently, ordinary listeners have insufficient cognitive strategies for consciously thinking about unfamiliar music while they listen, let alone deducing constructive rules that might contribute to their intellectual insight (if not necessarily to the aesthetic affect). In short, generative compositional rules tend to be completely analytical in nature.

Consider the rhythmic pattern of Figure 28a, from Berg’s *Wozzeck*. How many people who have heard this opera would consciously recognize that this pattern follows a strict rule? Of course, once pointed out, one can visually understand it, but in music listening, would our cognitions aurally map a rule of successive motivic diminution? The answer with respect to such constructivism is not clear.³⁴ Similar questions are even more pertinent with respect to the iterative additivity in the bottom notes of the sixteenth-note pattern of Figure 28b, from Stockhausen’s *Kruezspiel*, which lies even closer to a pure, generative rule than that seen in the Berg case. Other purely constructive instances include the superpositions and retrograded transformations of durational patterns that Messiaen promotes in his book *The Technique of My Musical Language* (1956), Carter’s precise use of metric modulation (Schiff, 1983), and the arithmetic rules governing *talea* in the music of Lutoslawski (Stucky, 1981).³⁵

Of course, many composers seem indifferent as to whether listeners hear such constructionisms. Indeed, the disconnect between twentieth-century compositional construction and music perception is one of the more puzzling aspects of contemporary culture, but it stems directly from nineteenth-century aesthetic beliefs that artists are prophets and therefore have a right

34. The relationship seen here is the simplest in this piece. The rest of the work is replete with inaudible rule-governed constructions, so common in the Continental music of this period.

35. Lutoslawski’s use of arithmetic techniques are often purely constructive and analytic, resistant to overlearning, and thus, as Stucky (1981) says, inaudible (though obviously affecting the texture). Contemporary music is full of such purely analytic rules.

Figure 28 consists of two musical examples, (a) and (b), illustrating constructionism.

Example (a) shows a musical phrase from Berg's *Wozzeck*. It begins with a melodic fragment consisting of a quarter note followed by a dotted quarter note. Above this, a bracket indicates "(successive diminutions of initial durational pattern)". The phrase continues with a series of notes, some of which are grouped with brackets and the number "3", suggesting triplets or similar rhythmic patterns.

Example (b) shows a musical phrase from Stockhausen's *Kruezspiel, I*, measures 7-13. It features a series of notes, some of which are grouped with brackets and the number "3", indicating triplets. Below the notes, a series of plus signs (+1, +2, +3, +4, +5, +6, +7, +8, +9, +10, +11) are placed, suggesting a sequence of additions or a specific rhythmic pattern. Above the notes, a bracket indicates "(successive additions to the lower note)".

Fig. 28. Constructionism. (a) Berg, *Wozzeck* (after Perle, 1980, p. 173); (b) Stockhausen, *Kruezspiel, I*, mm. 7–13 (♩ = 90) (discussed in Morgan, 1992, pp. 376–385).

to expect listeners to come to terms with what they create. Eschewing the relevance of ordinary perception and cognition altogether, several generations of contemporary music theorists, whose primary interests are to analyze music from the constructional point of view, have kept this essentially romantic view alive. Like their compositional counterparts, such theorists are completely unconcerned about whether their analyses make any cognitive sense in terms of ordinary music listening.

Other contemporary composers expect, nay, insist on conscious thought—even intense scrutiny—while listening to their music. Thus one rationale for modern analysis is to identify the specialized skills needed to comprehend this music. At its best, the aesthetic here resembles learning to play high-level chess or solving mathematical puzzles, and it should be said that music has always found room for this kind of ultra-connoisseurship (but perhaps never to the extent found in the mid-twentieth century). At its worst, the stipulated aesthetic is pretentious, even narcissistic.³⁶

From the scholarly view of music cognition, there are more important issues. Can we distinguish those cognitive rules of compositional construction that people can readily learn to hear from those analytic ones that are so difficult to perceive that people learn them only with considerable effort? What rule constructions do people apprehend “naturally” and which have to be “force-fed?” What kinds of rule making proceed automatically, what kinds require some training, and what kinds are totally generative (or analytic) and thus highly resistant to online processing regardless of the amount of training? Much work in music psychology is needed to answer such questions.

36. Of course, at the other extreme lies that large body of fatuous popular music that succeeds brilliantly in its perfect anti-intellectualism. Most serious music lovers find areas between these extremes the most rewarding.

Even in tonal music such questions crop up. Western music has typically celebrated contrapuntal music as a high point of compositional intellectualization. Yet perceptually oriented theorists have long pointed out that certain kinds of polyphonic textures are pure contrivances inasmuch as the contrapuntal devices used are not aurally perceived.³⁷

Although in tonal music working memory can process short retrogrades (and perhaps short retrograde inversions), as we saw in Figure 24, extended ones are very difficult to follow. Even with repeated study, long retrogrades remain cognitively obscure because they require the listener to treat all initial functions as terminal ones—and all terminal ones as initial—while also mapping medial functions in the reverse of how they were learned. Obviously a mathematical rule for deriving extended melodic retrogrades is unfeasible in online processing, even in the untransposed condition. Extended retrograde is taxing because it involves more “backward learning” than most ordinary listeners are willing to subject themselves to.

Experimental Issues and a Summary of the Musical Examples

The rules postulated to govern the musical examples are “natural” in the sense that they cover many musical styles from 1700 to the present. Presumably, then, any listener could learn to abstract these rules with sufficient exposure. The psychological question is, To what extent do listeners actually invoke cognitive rules to understand music? Although the explanations, descriptions, and analyses of the excerpts may be accurate (but by no means exhaustive), one may ask, Is rule invocation necessary (Equation 3)? Or is top-down, simple mapping all that listeners require (Equation 2)? Is music like language, a limited domain whose neural processing seems “prewired” for immutable bottom-up operations, or does music also rely on generalized similarities (be it instance, case, model, analogy, or category)? Does the cognition of music call for some hybrid rule system mediating between bottom-up and top-down processing? However we answer, the analytical observations presented here beg for experiment to confirm the extent to which listeners use cognitive rule-mapping. For if they do not construct and then apply rules, then we may relegate the preceding discussions of music to theoretical-analytical descriptions regarding merely different styles of composition. Many questions arise, to which we now turn.

THE QUESTION OF FORMAL SIMILARITY

We have seen formal similarity (A^0A^1) underscoring both complete sequence—whether contiguous (Figures 1, 3, and 4) or discontinuous (Fig-

37. Such schemes are, of course, discernible with study, and people can learn to hear them, but it is not what they naturally attend to.

ure 9a)—and partial sequence—whether both intervals diverge (Figure 8) or only one expands (Figures 5–7, 10–12). If, as argued, listeners invoke both rules of transformation ($^{0, 1}$) and top-down mappings of similarity (AA) in order to project the higher-level linear and triadic melodic sequences so familiar to tonal style, then they should also abstract and apply the same cognitive rules to unfamiliar styles. For that matter, listeners should apply rules even to highly abstract, noiselike signals, provided AA is present.³⁸ Indeed, novel contexts seem more likely to require abstract coding than highly familiar material. This is because rule invocation should evince maximum utility in unfamiliar situations, whereas mapping of mere instance may be all that's necessary in frequent, familiar contexts (Smith et al., 1989).

Listeners' sense of implied continuation in both familiar and unfamiliar as well as concrete and abstract cases could be tested via probe-tone technique. Krumhansl (1997) has already shown in tonal music that listeners do project higher-level continuations to some extent when formal similarity (AA) occurs. It remains to extend this work with respect to the rules and examples discussed in the present article. For inexperienced listeners, training might be required in tasks involving melodic sequences (whether contiguous or discontinuous, complete or partial), but this would not negate the independent reality of unconscious, abstract syntactical rules in music cognition if listeners then rule-mapped in unfamiliar contexts. One should also confirm whether the size of the higher-level interval of transposition makes any difference in the expected continuations of these cases (recall the discussion of reversal and similarity in Figure 2). One could then measure as well for degree of surprise (recall the denials and delays of sequential expectation discussed in earlier sections).

Of course, a range of relevant questions exist concerning the issue of musical similarity. How does ornamental variation (symbolized by the superscript letters of the As) affect higher-level projection? Is there a difference in reaction time between recognition of exact sameness (A^0A^0) versus recognition of various degrees of similarity (A^0A^1)? Is a pattern followed by diminution less likely to be heard as A^0A^1 (recall Figure 15a) than one followed by augmentation (re Figure 15b)? Trials might also vary type of parametric change to see whether some parameters contribute more to recognition of similarity than others. Is, for example, durational similarity more important than the variables of melodic similarity (whether of contour, interval, or scale-step function)?³⁹

38. Abstract formal similarity (AA) will ensure the hypothesis that the stylistically different materials to be tested share the property of being governed by the same rule.

39. By parametric change I do not refer here to augmentation or ongoing diminution because in these procedures the durational ratios of the overall pattern remain the same. Rather I refer to durational change that is differentiated note by note so that no previous temporal property remains constant.

THE QUESTION OF FORMAL DISSIMILARITY

How much parametric change is necessary before listeners no longer hear similarity and thus no longer melodically project full or partial sequence according to abstract rule? Stated operationally, when do listeners cognitively convert A^0A^1 to AB (recall Table 1)? Is there a threshold of formal differentiation that prevents the perception of higher-level implications? And how does the functional location of ornamental variation affect the perception of higher-level, sequential expectation (recall the discussion of Figure 26)? Given an AA and AB that are similarly closed, is there less sense of expected higher-level continuity after an AB ?

At what point does a shift in parametric differentiation affect the degree of closure? On the basis of the terminal segmentations that we have seen (e.g., Figures 11b and 15b), one guesses that a change in subset at the beginning of an A^1 would be more disruptive to listeners' projections of higher-level sequential transpositions than a change at the end (and what about the ornamentation of a subset in the middle of such a motivic form?). As one varies the amount of parametric change across the formal spectrum, are there primacy or recency affects? Along with length of pattern, one expects, for example, that the degree of parametric variation would affect the perception of short inversions and retrogrades (not to mention short retrograde inversions).

THE QUESTION OF PARTIAL SEQUENCE, REGISTRAL RETURN, AND STREAMING

Concerning partial sequences, which often involve both exact and near registral return (*aba* occurs throughout in Figures 5–7 and 10–12), we need experiments to determine whether cognitive rules applied to either expanding or contracting intervals can be distinguished from the phenomenon of voice streaming. Streaming may be hardwired (Bregman, 1990), but then registral return may be as well. If both are, then the question is whether rule-mapping is even necessary to perceive partial sequence. Recall in Figure 5a that registral return was exact (on F), but the alternating tones (A, C, F²) exceeded the definition of return (according to the implication-realization model),⁴⁰ yet the ear clearly hears the triadic connection between these tones (presumably according to streaming). So is the invocation of cognitive rule in these cases of partial sequence necessary?⁴¹

40. In the implication-realization model, registral return (*aba*) is a constrained subset of the psychological phenomenon of streaming, where it occurs only when discontinuous pitches create an interval no greater than a major second.

41. That invoking cognitive rule for partial sequence appears redundant in the light of both registral return and streaming is not a *prima facie* argument against rule-making because streaming itself may result from cognitive rule. Moreover, redundancy is an inherent

One place to test this would be in contexts of formal differentiation (*AB*). We would attempt to determine in forms of *AB* whether listeners hear registral return in one voice (a . . . a) and streaming via nonregistral return in another (b, c, d). Probe-tone experiments could determine this. If both streaming and registral return are innately bottom-up activities, then listeners should still perceive both in *AB* conditions where higher-level cognitive rules are less relevant.

THE QUESTION OF FORMAL DIFFERENTIATION

Of course, just as rule-mapping can occur in formal similarity (A^0A^1), so it can occur in formal differentiation (*AB*). Recall the continuing intervallic expansion in the differentiated forms of Figure 11a or the increasing and then decreasing dissonance in Figure 20d, where the motivic melodic forms are generally quite dissimilar. In sum, the extent to which whether perceivers invoke rules in both *AA* and *AB* musical forms needs to be tested.

To see whether listeners rely on analogy to stored instance or on rule-mapping to comprehend patterns is of crucial importance. Because the majority of examples herein have occurred in contexts of formal similarity, the question of whether rule-mapping is necessary again arises. Formal similarity seems to be allied with comparison by analogy. In a musical sequence, for instance, transposition could emerge as a result of comparing variant (A^1) with model (A^0). And if a second transposed variant (A^2) occurred, then the comparative task looks very much like “ A^0 is to A^1 as A^1 is to A^2 ” (rational model: “A is to B as B is to C”).

To examine whether analogy explains a listener’s response to sequence better than rule-mapping, we must test in the *AB* condition. Specifically, we would set up an implication in formal similarity (*AA*) and then realize it in a differentiated form (*B*). This should not be too hard to model, as *AAB* contexts in music are extremely common. Figure 29 illustrates the possibility. If through probe tones listeners hear the C# in the first B^0 (m. 3) as expected and thus unsurprising, then the expectation generated by the sequence is not caused by formal similarity (*AA*) or analogical thinking but by rule-mapping.⁴²

property of communication systems. Thus it may be that in other contexts all three—streaming, registral return, and rule—are necessary to arrive at the contextual meaning. The melody of Figure 17, for example, exhibits near registral return while the contracting and expanding of intervals seem rule governed.

42. There are many aspects about analogies that I cannot discuss here. Analogies do not generalize nearly as specifically as abstract rules and thus cannot deal with completely unfamiliar material as precisely. Moreover, analogies are initially retrospective in nature, and this does not fit well with the notion of expectation, which is so important to music. Finally, analogies seem to be consciously invoked when problems arise, whereas the online processing of music is largely unconscious and unproblematic.

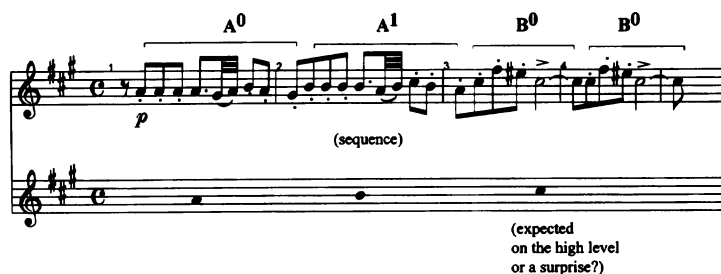


Fig. 29. Testing for the effects of formal differentiation on the perception of sequence. Tchaikovsky, *Swan Lake*, Act 2, Danse des cygnes (no. 13), part 4, mm. 2–4 (Allegro moderato).

THE QUESTION OF MEASURING AFFECT

One expects with similarly related input that iterated applications of rule will have a priming effect, facilitating access to the rule and thus improving the listener's reaction time in projecting higher-level continuations. Accordingly, an interrupting surprise of rule activation accompanying formal similarity (A^0A^1) would presumably be more affective than one with formal differentiation (AB) because stylistic similarity generates prospective continuation based on learned probability more strongly. (It seems likely that this is what drove Strauss to compose melodies like those found in Figure 6a–e, each new passage varying a previously established partial sequence.)

Formal contexts of AA probably enhance listeners' expectations of continuation and their reaction to its interruption. Consequently, repetition per se should help refine rule-mapping, thereby improving its accuracy. As exposure to formal repetition and hence similarity increases, listeners probably begin to convert the materials of analogous cases into rule variables. With practice this updating should in turn transform previously learned context-specific rules into more powerful abstractions. All this would influence the degree of affect felt upon denial.

Thus similarity-based learning should facilitate both the accessibility of rule and its transferability to higher-level domains. Experiments involving musical training in generalized rules should show these benefits, not only in different stylistic contexts but also in analogously derived applications in other contrasting styles. One would expect, for example, that listeners carefully trained to hear tonal sequences (say, in the music of Bach or Handel) would be better equipped to hear such transforms in atonal music (say, in Skryabin or early Schoenberg). Likewise, one would think that with sufficient training listeners could be taught to hear not only relatively brief types of mirroring (e.g., the various styles seen in Figures 22a–c and 24a–b) but also more subtle cases of it (e.g., Figure 25b).

As for transferability of applications across parameters, it may even be the case that listeners adequately trained in hearing subtractive rhythmic groups, such as Berg used (Figure 28a), could readily transfer that abstraction to successively subtractive forms, such as those found in Beethoven's music (Figure 27). Experiments based on these possibilities would train participants to play an intellectual game—to find the math or the logic in the music—and then test to see whether they could map the learned cognitive rules to different parametric variables.⁴³

THE QUESTION OF MULTIPLE RULES

In many cases, rule-mapping may be more difficult and thus less accurate if listeners find that more than one rule is simultaneously required. Further, the need to invoke two or more rules would presumably slow down listeners' reaction time in projecting higher-level continuation or in assessing retrospective relationships. The hypotheses of multiple rules reducing both accuracy and reaction time are generally reasonable (Smith et al., 1993) but need to be tested in music. For example, patterns where both melodic and durational rules are highly coordinated might not show either a loss of accuracy or a slower reaction time. However, melodies such as the purely additive case seen in Figure 16a or the shorter durations and concomitantly increased sense of pace in Figure 16b may not show these effects. This is because in these cases experienced listeners learn that a single correlated rule governs the melodic-durational patterning.

A multirule pattern like Figure 18b, however, which (1) adds melodic tones, (2) systematically expands intervallic size, and generally (3) decreases the duration of melodic tones (which increases the pace) is more complex. Complexity is even greater in the "wedge" pattern of Figure 20c, which employs (1) a systematic increase in dissonance, (2) increasing density of texture, (3) increasingly smaller durations (and hence a systematic increase in pace), and (4) simultaneous inversion in outside pairs of voices (which may not be perceived at all and thus exist purely as a compositional device). In three- or four-rule contexts like these, both reaction time and accuracy in rule-projecting the continuation of all the parametric particulars will probably suffer.

It is also the case, one believes, that some parametric rules are more easily accessible than others. Figure 16c, for instance, also displays a com-

43. By using functional magnetic resonance imaging, Dehaene, Spelke, Pined, Stanesco, and Tsivkin (1999) show evidence that indicates that people perform simple addition and multiplication in the left frontal lobe of the brain (the area thought to store language memory), whereas simple subtraction (and possibly division) seems to take place in an area dedicated to scaled spatial tasks (the parietal lobes). In light of this, it would be interesting to know whether musical augmentation, which involves multiplication (recall Figure 14a–b), is processed in an area different from the area where musical diminution, which relies on division (recall Figure 15a), is processed.

plex sets of rules, but its general direction toward adding tones of faster durations while increasing the pace is clearly more recognizable than, say, simultaneous inversion in contrary pairs of voices.

Patterns where one rule is nested in another, as in Figures 7a and 7c (partial sequences nested in full sequences) and 15b (partial sequence embedded in augmentation), probably endure no such loss in reaction time or accuracy, particularly because in both these cases the extraopus stylistic material is very familiar (made so by the intraopus stylistic repetition) and thus easily “chunked” in working memory.⁴⁴ For that matter, listeners probably also find it easy to follow alternating sequences in opposite directions, as seen in Figure 9a.

Conclusion

In the absence of rules, listeners project expectations by building analog models based on “pure instance” (Smith et al., 1993). We first tend simply to map varied repetitions of a given instance from the top down, without attempting rule construction. Only after repeated exposure do listeners begin to abstract cognitive rules. We thus arrive at deductive formulations through inductive learning (the amount of learning necessarily varying with the nature of the pattern). This is why a lot of cognitive rule-making goes hand in hand with stylistic repetition. Indeed, rules are probably intrinsic to style’s infrastructure. As we have seen, in music many kinds of cognitive rules always seem to occur in contexts of formal similarity (AA).

Prospective cognitive rules involving formal differentiation (AB) probably take longer to learn than those involving AA. This is because they are harder to formulate. When listeners encounter differentiation, they have to devote most of their attention to processing details on the manifest level before constructing an abstract rule. Retrospective cognitive rules involving AB are bound to the surface materials. Proscriptive or retrospective rules that emanate from composition alone, whether in AA or AB forms, are merely constructive (and by inference analytic). Such “pure rules” require much top-down training in order to be cognitively applied in online processing (if at all).

What characterizes music cognition is not the variety of cognitive strategies that humans have for projecting expectations. Rather it is that so many of these—recognizing process or reversal, mapping intraopus or extraopus style, deducing sequence, partial sequence, mirrored symmetries,

44. But as Jones (1981, pp. 496–497) reminds us, hierarchical sequences are not necessarily easier to remember than nonhierarchical ones. It all depends on the nature of the pattern.

and so forth—are all based in one way or another on one ability, that of mapping *abstractions* of similarity or dissimilarity. This is so whether one is dealing with the input of scaled parametric materials ($a + a$, as in melody, duration, harmony, etc.), projecting formal reproductions or rule-mapping variations (A^0A^0 , A^0A^1), tracking ornamental change for content's sake (A^aA^b), or automatically linking the same kind of learned, extraopus cognitive rule to a variety of novel input.

Our comprehension of the world, whether it is music or anything else, is powerful precisely because high-level rules cover a huge variety of phenomena, including, as the copious musical examples in this article show, the richness of many different musical styles. This is true whether such abstractions cover parameters, where $a + a \rightarrow a$ (where each a is an element or a primitive), or forms, where $A + A \rightarrow A$ (where each A is a closed group), or rules, where $^0 + ^1 \rightarrow ^2$ (where 1 and 2 represent either additivity, subtractivity, divisibility, multiplicativity, or some combination thereof applied to a variable).

One concludes that humans are similarity automatons regardless of domain, level, or operation. Such cognitive dependence of so much on so little stems directly from evolution, where the ubiquitous complexity found in life is comprehensible only because the high-level abstract rules governing our unconscious processing are so very simple—and for that reason so very powerful.⁴⁵

References

- Arnheim, R. (1969). *Visual thinking*. Berkeley: University of California Press.
- Bartlett, J. C., & Dowling, W. J. (1980). Recognition of transposed melodies: A key-distance effect in developmental perspective. *Journal of Experimental Psychology: Human Perception and Performance*, 6, 501-515.
- Berry, W. (1966). *Form in music*. Englewood Cliffs, NJ: Prentice-Hall.
- Boltz, M., & Jones, M. R. (1986). Does rule recursion make melodies easier to reproduce? If not, what does? *Cognitive Psychology*, 18, 389-431.
- Boltz, M., Marshburn, E., Jones, J. R., & Johnson, W. W. (1985). Serial-pattern structure and temporal-order recognition. *Perception & Psychophysics*, 1985, 209-217.
- Bowen, J. (1996). Tempo, duration, and flexibility: Techniques in the analysis of performance. *Journal of Musicological Research*, 16, 1-47.
- Bregman, A. S. (1990). *Auditory scene analysis*. Cambridge, MA: MIT Press.
- Bruce V. (1991). *Face recognition*. Hillsdale, NY: Erlbaum.
- Clarke, E. F. (1989). The perception of expressive timing in music. *Psychological Research*, 51, 2-9.
- Cooper, G., & Meyer, L. B. (1960). *The rhythmic structure of music*. Chicago: University of Chicago Press.
- Crist, W., DeLone, R., Kliever, V., Rowell, L., & Thomson, W. (1982). *Materials and structure of music*. Englewood Cliffs, NJ: Prentice-Hall.

45. I am grateful to Caroline Palmer and Denisa N. Borges for their help with this article.

- Dehaene, S., Spelke, E., Pinel, P., Stanescu, R., & Tsivkin, S. (1999). Sources of mathematical thinking: Behavioral and brain-imagine evidence. *Science*, 284, 970–974.
- Deutsch, D. (1999). The processing of pitch combinations. In D. Deutsch (Ed.), *The psychology of music* (2nd ed., pp. 349–411). San Diego: Academic Press.
- Dowling, W. J. (1972). Recognition of melodic transformations: Inversion, retrograde, and retrograde inversion. *Perception & Psychophysics*, 12, 417–421.
- Dowling, W. J. (1978). Scale and contour: Two components of a theory of memory for melodies. *Psychological Review*, 85, 341–354.
- Dowling, W. J. (1991). Tonal strength and melody recognition after long and short delays. *Perception & Psychophysics*, 50, 305–313.
- Dowling, W. J. (1994). Melodic contour in hearing and remembering melodies. In R. Aiello (Ed.), *Musical perceptions* (pp. 173–190). New York: Oxford University Press.
- Dowling, W. J., & Bartlett, J. C. (1981). The importance of interval information in long-term memory for melodies. *Psychomusicology*, 1, 30–49.
- Dowling, W. J., & Fujitani, D. S. (1971). Contour, interval, and pitch recognition in memory for melodies. *Journal of the Acoustical Society of America*, 49, 524–531.
- Dowling, W. J., Kwak, S., & Andrews, M. (1995). The time course of recognition of novel melodies. *Perception & Psychophysics*, 57, 136–149.
- Edworthy, (1985). Interval and contour in melody processing. *Music Perception*, 2, 375–388.
- Epstein, D. (1985). Tempo relations: A cross-cultural study. *Music Theory Spectrum*, 7, 34–71.
- Epstein, D. (1995). *Shaping time: Music, the brain, and performance*. New York: Schirmer.
- Feldman, J., Epstein, D., & Richards, W. (1992). Force dynamics of tempo change in music. *Music Perception*, 10, 185–203.
- Francès, R. (1988). *The perception of music* (W. J. Dowling, Trans.). Hillsdale, NJ: Erlbaum. (Original work published 1957)
- Gattis, M., & Holyoak, K. J. (1996). Mapping conceptual to spatial relations in visual reasoning. *Journal of Experimental Psychology; Learning, Memory, and Cognition*, 22, 231–239.
- Gjerdingen, R. (1988). *A classic turn of phrase: Music and the psychology of convention*. Philadelphia: University of Pennsylvania Press.
- Goldstone, R. L. (1993). Feature distribution and biased estimation of visual displays. *Journal of Experimental Psychology: Human Perception and Performance*, 19, 564–579.
- Green, D. (1965). *Form in tonal music*. New York: Holt, Rinehart and Winston.
- Halpern, A. R., Bartlett, J. C., & Dowling, W. J. (1998). Perception of mode, rhythm, and contour in unfamiliar melodies: Effects of age and experience. *Music Perception*, 15, 335–355.
- Hofstadter, D. R. (1979). *Gödel, Escher, Bach: An eternal golden braid*. New York: Basic Books.
- Hudson, R. (1994). *Stolen time: The history of tempo rubato*. Oxford: Clarendon Press.
- Jones, M. R. (1974). Cognitive representations of serial patterns. In B. H. Kantowitz (Ed.), *Human information processing: Tutorials in performance and cognition* (pp. 187–229). Hillsdale, NJ: Erlbaum.
- Jones, M. R. (1981). A tutorial on some issues and methods in serial pattern research. *Perception & Psychophysics*, 30, 492–504.
- Jones, M. R. (1987). Dynamic pattern structure in music: Recent theory and research. *Perception and Psychophysics*, 41, 621–634.
- Jones, M. R. (1990). Learning and the development of expectancies: An interactionist approach. *Psychomusicology*, 9, 193–228.
- Jones, M. R., & Ralston, J. T. (1991). Some influences of accent structure on melody recognition. *Memory and Cognition*, 19, 8–20.
- Kihlstrom, J. F. (1987). The cognitive unconscious. *Science*, 237, 1445–1452.
- Krumhansl, C. L. (1997). Effects of perceptual organization and musical form on melodic expectancies. In M. Leman (Ed.), *Music, gestalt, and computing* (pp. 294–320). Berlin: Springer.

- Krumhansl, C. L., Sandell, G. J., & Sargeant, D. C. (1987). The perception of tone hierarchies and mirror forms in twelve-tone serial music. *Music Perception*, 4, 153–184.
- Marcus, G. F., Vijayan, S., Rao, S. Bandi, & Vishton, P. M. (1999). Rule learning by seven-month-old infants. *Science*, 283, 77–80.
- Matthay, T. (1913). *Musical interpretation*. Boston: Boston Music.
- McEwen, J. B. (1928). *Tempo rubato or time-variation in musical performance*. Oxford: Oxford University Press.
- Medin, D. L., Goldstone, R. L., & Gentner, D. (1993). Respects for similarity. *Psychological Review*, 100, 254–278.
- Medin, D. L., Goldstone, R. L., & Markman, A. B. (1995). Comparison and choice: Relations between similarity processes and decision processes. *Psychonomic Bulletin and Review*, 2, 1–19.
- Messiaen, O. (1956). *The technique of my musical language* (J. Salterfield, Trans.). Paris: Alphonse Leduc.
- Meyer, L. B. (1956). *Emotion and meaning in music*. Chicago: University of Chicago Press.
- Meyer, L. B. (1973). *Explaining music*. Berkeley: University of California Press.
- Meyer, L. B. (1989). *Style and music*. Philadelphia: University of Pennsylvania Press.
- Morgan, R. P. (1992). *Anthology of twentieth-century music*. New York: W. W. Norton.
- Narmour, E. (1977). *Beyond Schenkerism: The need for alternatives in music analysis*. Chicago: University of Chicago Press.
- Narmour, E. (1988). Melodic structuring of harmonic dissonance: Chopin's contribution to the development of harmony. In J. Samson (Ed.), *Chopin studies* (pp. 77–114). Cambridge: Cambridge University Press.
- Narmour, E. (1990). *The analysis and cognition of basic melodic structures: The implication-realization model*. Chicago: University of Chicago Press.
- Narmour, E. (1991). The top-down and bottom-up systems of musical implication: Building on Meyer's theory of emotional syntax. *Music Perception*, 9, 1–26.
- Narmour, E. (1992). *The analysis and cognition of melodic complexity: The implication-realization model*. Chicago: University of Chicago Press, 1992.
- Narmour, E. (1996). Analyzing form and measuring perceptual content in Mozart's piano sonata K. 282: A new theory of parametric analogues. *Music Perception*, 13, 265–318.
- Narmour, E. (1999). Hierarchical expectation and musical style. In D. Deutsch (Ed.), *The psychology of music* (2nd ed., pp. 441–472). New York: Academic Press.
- Nisbett, R. E. (1993). Reasoning, abstraction, and the prejudices of 20th century psychology. In R. E. Nisbett (Ed.), *Rules for reasoning* (pp. 1–12). Hillsdale, NJ: Erlbaum.
- Palmer, C. (1989). Mapping musical thought to musical performance. *Journal of Experimental Psychology: Human Perception and Performance*, 15, 331–346.
- Pazzani, M. J. (1991). Influence of prior knowledge on concept acquisition: Experimental and computational results. *Journal of Experimental Psychology: Learning, Memory, and Cognition*, 17, 416–432.
- Perle, G. (1977). *Twelve-tone tonality*. Berkeley: University of California Press.
- Perle, G. (1980). *The operas of Alban Berg* (Vol. 1). Berkeley: University of California Press.
- Pinker, S. (1999). Out of the minds of babes. *Science*, 283, 40–41.
- Randel, D. M. (1986). *The new Harvard dictionary of music*. Cambridge, MA: Belknap Press.
- Repp, B. H. (1992). A constraint on the expressive timing of a melodic gesture: Evidence from performance and aesthetic judgment. *Music Perception*, 10, 221–241.
- Repp, B. H. (1994). Relational invariance of expressive microstructure across global tempo changes in music performance: An exploratory study. *Psychological Research*, 56, 269–284.
- Repp, B. H. (1997a). The aesthetic quality of a quantitatively average music performance: Two preliminary experiments. *Music Perception*, 14, 419–444.
- Repp, B. H. (1997b). Expressive timing in a Debussy prelude: A comparison of student and expert pianists. *Musicae Scientiae*, 1, 257–266.

- Restle, F. (1970). Theory of serial pattern learning: Structural trees. *Journal of Experimental Psychology*, 92, 385–390.
- Rosner, B., & Meyer, L. B. (1986). The perceptual roles of melodic process, contour, and form. *Music Perception*, 4, 1–39.
- Sadler, D. D., & Shoben, E. J. (1993). Context effects on semantic domains as seen in analogy solution. *Journal of Experimental Psychology: Learning, Memory, and Cognition*, 19, 128–147.
- Salzer, F. (1962). *Structural hearing*. New York: Dover.
- Schenker, H. (1954). *Harmony*. Chicago: University of Chicago Press.
- Schiff, D. (1983). *The music of Elliott Carter*. London: Eulenburg.
- Schoenberg, A. (1975). *Style and idea* (L. Stein, Ed.). London: Faber and Faber. (Original work published 1950)
- Schuller, G. (1997). *The compleat conductor*. New York: Oxford University Press.
- Selfridge-Field, E. (1998). Conceptual and representational issues in melodic comparison. In W. B. Hewlett & E. Selfridge-Field (Eds.), *Melodic similarity: Concepts, procedures, and applications* (Computing in Musicology 11). Stanford: Center for Computer Assisted Research in the Humanities.
- Simon, H. A. (1972). Complexity and the representation of patterned sequences of symbols. *Psychological Review*, 79, 369–382.
- Slobodkin, L. B. (1992). *Simplicity and complexity in games of the intellect*. Cambridge, MA: Harvard University Press.
- Smith, E. E., Langston, E., & Nisbett, R. E. (1993). The case for rules in reasoning. In R. E. Nisbett (Ed.), *Rules for reasoning* (pp. 361–397). Hillsdale, NJ: Erlbaum.
- Solso, R. L. (1994). *Cognition and the visual arts*. Cambridge, MA: MIT Press.
- Stucky, S. (1981). *Lutoslawski and his music*. Cambridge: Cambridge University Press.
- Sundberg, J., & Verrillo, V. (1980). On the anatomy of the ritard: A study of timing in music. *Journal of the Acoustical Society of America*, 68, 772–779.
- Takeuchi, A. H., & Hulse, S. H. (1992). Key-distance effects in melody recognition reexamined. *Music Perception*, 10, 1–23.
- Tenney J., with Polansky, L. (1980). *Meta + hodos—a phenomenology of 20th century music and an approach to the study of form*. New Orleans: Tulane University Press.
- Toch, E. (1977). *The shaping forces of music*. New York: Dover Publications.
- Todd, N. P. (1985). A model of expressive timing in tonal music. *Music Perception*, 3, 33–58.
- Todd, N. P. (1995). The kinematics of musical expression. *Journal of the Acoustical Society of America*, 97, 1940–1949.
- Tovey, D. F. (1965). *Essays in musical analysis* (Vol. I). Oxford: Oxford University Press.
- Tversky, B. (1977). Features of similarity. *Psychological Review*, 84, 327–352.
- Uttal, W. R. (1988). *On seeing forms*. Hillsdale, NJ: Erlbaum.
- van Egmond, R., & Povel, D.-J. (1996). Perceived similarity of exact and inexact transpositions. *Acta Psychologica*, 92, 283–295.
- van Egmond, R., Povel, D.-J., & Maris, E. (1996). The influence of height and key on the perceptual similarity of transposed melodies. *Perception & Psychophysics*, 58, 1252–1259.
- Vos, P. G., van Assen, M., & Franek, M. (1997). Perceived tempo change is dependent on base tempo and direction of change: Evidence for a generalized version of Schulze's (1978) internal beat model. *Psychological Research*, 59, 240–247.
- Wagner, R. (1897). *On conducting* (E. Dannreuther, Trans.). London: W. Reeves. (Original work published 1869)
- Watkins, A. J. (1985). Scale, key, and contour in the discrimination of tuned and mistuned approximations to melody. *Perception & Psychophysics*, 37, 275–285.
- Willats, J. (1998). *Art and representation*. Princeton, NJ: Princeton University Press.
- Winters, T. (1986). *Additive and repetitive techniques in the experimental works of Charles Ives*. Unpublished doctoral dissertation, University of Pennsylvania.
- Wisniewski, E. J., & Medin, D. L. (1994). On the interaction of theory and data in concept learning. *Cognitive Science*, 18, 221–281.