META ⁄ HODOS*

A Phenomenology of 20th-Century Musical Materials
and an Approach to the Study of Form

and

META Meta ⁄ Hodos

James Tenney

F 45'88

* "meth' od, n. [F. *methode*, fr. L. *methodus*, fr. Gr. *methodos*, method, investigation following after, fr. *meta* after + *hodos* way.] . . ."

Second edition, 1988

Originally published by the Inter-American Institute for Musical Research, Tulane University, New Orleans, 1964 (*META+HODOS*), and the *Journal of Experimental Aesthetics* 1.1, 1977 ("META Meta+Hodos").

editor: Larry Polansky
design and production: Carter Scholz
copy editor: Lauren A. Pratt
typing: Claudia Hasegawa

Special thanks to Carol Lennox, Marcia Williams, and the Mills College Computer Learning Studio.

FROG PEAK MUSIC (A Composers' Collective)

Contents

Publisher's Introduction

Meta+Hodos was originally written by James Tenney as his Master's Thesis at the University of Illinois, at Champaign-Urbana, in 1961. It was published in a very limited edition by Gilbert Chase some years later, but has been nearly impossible to obtain since its creation. Yet it has had a wide and powerful impact on music theory and composition in the past 25 years, to a degree greatly disproportionate to its availability. *META Meta+Hodos*, written in 1975, was first published in the *Journal of Experimental Aesthetics* (Volume 1, Number 1, 1977). The present Frog Peak Music edition of *Meta+Hodos* and *META Meta+Hodos* marks an attempt to make these seminal theoretical documents available to a larger community of artists.

This second edition includes corrections and revisions by the author.

Larry Polansky
Oakland, 1988

META/HODOS

June 1961

Section I.

"A good description of a phenomenon may by itself rule out a number of theories and indicate definite features which a true theory must possess. We call this kind of observation 'phenomenology', a word which means. . .as naïve and full a description of direct experience as possible."

Kurt Koffka, *Principles of Gestalt Psychology*, p. 73.

". . .one must be convinced of the infallibility of one's own fantasy and one must believe in one's own inspiration. Nevertheless, the desire for a conscious control of the new means and forms will arise in every artist's mind, and he will wish to know consciously the laws and rules which govern the forms which he has conceived 'as in a dream'. Strongly convincing as this dream may have been, the conviction that these new sounds obey the laws of nature and our manner of thinking. . .forces the composer along the road of exploration."

Arnold Schoenberg, *Style and Idea*, p. 218.

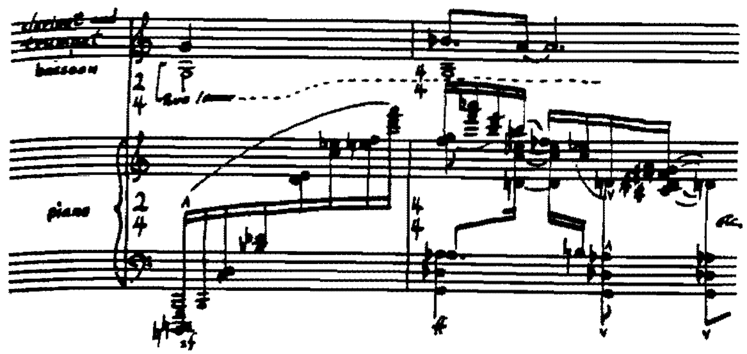
"The first step in the direction of beauty is to understand the frame and scope of the imagination, to comprehend the act itself of esthetic apprehension."

James Joyce, *A Portrait of the Artist as a Young Man*, p. 208.

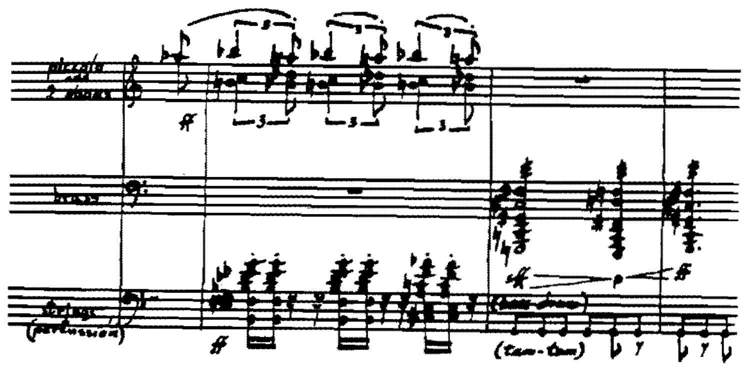
Section I.
The New Musical Materials.

The increased aural complexity of much of the music of the 20th century is such an evident characteristic that it should need no demonstration. Nevertheless, an examination of the many factors which produce this complexity, and of some of its effects in our perception of the music, will be necessary before we can hope to describe the musical materials in a really meaningful way. The complexity is not merely of structure, but also of substance. That is, it is not simply the result of a new arrangement of traditional materials or elements (I shall use the word *element* in this book in the sense of "part" or "portion," rather than "aspect" or "factor"). The elements themselves have changed, and the changes affect not only the musical structure, but our way of listening to the music as well. And the problems which arise from this seem to go beyond the mere question of the amount of time required for the ear and mind to assimilate the novelties of a new style until they no longer have what Schoenberg once described as a "sense-interrupting effect." Time has given us some degree of familiarity with even the most advanced musical achievements of the early 20th century, and yet our descriptive and analytical approaches to this music are still belabored with negatives—"atonal," "athematic," etc.—which tell us what the music is *not*, rather than what it is. The narrowness of the traditional musical concepts is manifested by this very negativism, and by the fact that many significant works of this earlier period are too often relegated to the realm of "exceptions," "deviations," or "interesting experiments." And the disparity between the traditional concepts and the actual musical "object" becomes even greater with the more recent (non-instrumental) electronic and tape-music. But even here, the problem is not really one of a lack of familiarity, but of a nearly complete hiatus between music theory and musical practice. Thus, even when the novelties of the various styles and techniques of 20th-century music have become thoroughly familiar, certain "complexities" will still remain outside of our present conceptual framework, and it is clear that this conceptual framework is in need of expansion.

I have said that the materials of the music have changed, and this is to be seen in countless examples in which the primary musical ideas are highly complex sound-configurations whose basic



Example 1. Charles Ives, "Scherzo" (p. 16).



Example 2. Anton Webern, op. 6, #2 (mm. 17-19).

Example 3. Béla Bartók, Sonate (piano), (p.18).

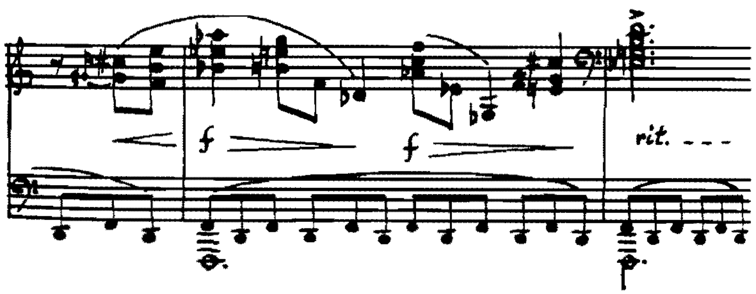
elements are themselves more or less complex structures rather than single tones. Typical configurations of this kind are shown in examples 1 to 3. Such elemental sound-structures occur in a great variety of forms, with respect to both their vertical structure and their changes in time. I shall examine them first from the standpoint of their vertical structure, with particular attention to elements in which the vertical structure is a more noticeable characteristic than any temporal form they may have.

The clearest examples of such complex sound-elements are tone-clusters and other highly dense and dissonant chords, as in these first three examples—sound-structures which seem relatively "opaque" to the ear. Such chords cannot usually be analyzed by the ear into constituent tones, and I think they are not intended to be so analyzed. They are seldom subject to harmonic orientation, because one's perception of pitch in these dense sound-complexes is limited, at best, to the pitch of their highest or lowest tones, or to a mean pitch-level, when no more than the approximate range and register of the chords can be recognized. Their similarity to percussive sounds is very close, and it is significant that the use of such complex sound-elements coincides historically with an increasing exploitation of the percussion instruments of the orchestra, and that they are frequently to be found in music of an intentionally "rhythmic" or motoric character, such as the Bartók sonata from which example 3 is taken. Such chords represent, in fact, a kind of bridge between more traditional harmonic structures and purely percussive sounds and noises, and it would be difficult to find any clear-cut line of distinction between any two of these three types of sound-elements. They are distinguished from each other only in the relative difficulties they present to the ear's power of pitch analysis, and thus in their relative specificity of pitch-definition, and in the possibility of harmonic orientation which depends on such pitch-definition. The percussion battery itself includes both instruments of definite pitch and ones of indefinite pitch, and the sounds produced by the latter instruments are nothing more than "tone-clusters" of a higher degree of complexity.

There is thus a continuous "spectrum" of composite sound-elements, ranging from simple chords whose constituent tones can be analyzed by the ear—through more complex and opaque sounds whose pitch-characteristics are more or less indefinite, or only partially perceptible—to sounds without any definite pitch, which we characterize as noise. But in spite of the breadth of this spectrum,

examples can be found of the use of each of these three types of composite sounds as essentially *irreducible elements* of musical ideas—examples in which such sound-complexes are substantially *equivalent* to single tones.

One manifestation of the gradual use of more and more complex sound-units in place of single tones is to be seen in the expansion of the very concept of "melodic line" by way of various kinds of doublings. This concept had already been somewhat complicated in pre-20th-century music by the frequent doublings in thirds and sixths, and in the late 19th century by the use of parallel seventh and ninth chords. These devices were intended to enrich the sonority of a single melodic line, without adding any really independent lines to the texture, and the intervals and chords so used can fairly be said to be equivalent to single tones, with respect to most of the formal functions. But by about 1910, these devices had been considerably extended to include not only other, more dissonant intervals and chords, but also more complex "doublings" in which the intervals change in the course of a single line, or in which the number of tones in each element is varied from one to the next—and often both types of variation are employed within the same line, as in example 4.



Example 4. Arnold Schoenberg, op.11, #2 (p.7).

There was a time when theorists could refer to noises as "non-musical sounds," and this attitude still exists to some extent. But it is clearly unrealistic to make such a distinction now, in the light of musical developments in the 20th century. The elemental building-materials of this music are no longer limited to "musical" tones, but may include other, more complex sounds, which in an earlier music would have seldom functioned as elements, if they occurred at all. *The substance and material of this music is sound*—this definition is inescapable—and it is of secondary importance whether this material is in the form of a tone with clearly defined pitch, or of the highly complex and indefinitely-pitched sound of a cymbal. *Any sound might occur* at some point in a piece of music, with a function there that is virtually independent of the constitution or structure of the sound itself, being determined instead by the larger musical context in which it occurs. Once this is acknowledged, it becomes evident that the first requisite of an expanded conceptual framework for the music of our time will be a principle of *equivalence*, by which recognition is made of the equal potentiality of any sound being used as a basic element in a musical idea.

The full implications of this principle will become more clear in the course of the book, but here it may be noted that there is a close parallel to this idea of equivalence in Schoenberg's arguments about consonance and dissonance, and an examination of this parallel may help to elucidate the idea being presented here. In *Style and Idea* [1] Schoenberg says:

> "What distinguishes dissonances from consonances is not a greater or lesser degree of beauty, but a greater or lesser degree of *comprehensibility*. In my *Harmonielehre* I presented the theory that dissonant tones appear later among the overtones, for which reason the ear is less intimately acquainted with them. This phenomenon does not justify such sharply contradictory terms as concord and discord. Closer acquaintance with the more remote consonances—the dissonances, that is—gradually eliminated the difficulty of comprehension and finally admitted not only the emancipation of dominant and other seventh chords, diminished sevenths and augmented triads, but also the emancipation of Wagner's, Strauss's, Moussorgsky's, Debussy's, Mahler's, Puccini's, and Reger's more remote dissonances.
>
> "The term *emancipation of the dissonance* refers to its comprehensibility, which is considered equivalent to the consonance's comprehensibility. A style based on this premise treats dissonances like consonances and renounces a tonal center. . ."

Now there is an apparent inconsistency in this argument—that is, if we understand the word "equivalent" (in the second paragraph) in an unnecessarily restricted way—because he has not established a real equivalence of comprehensibility as such, but simply a *relativity* of consonance and dissonance, and a lack of any clear-cut distinction or opposition between them. I suggest that he means a different sort of equivalence, and one which is analogous to the *principle of equivalence* I am proposing here. It is a functional equivalence that Schoenberg is describing, which postulates the equal potentiality of both consonances and dissonances being used as material in the musical texture—in spite of their differences with respect to "comprehensibility." In other words, the relative consonance or dissonance of a sound is no longer considered to be a functionally relevant characteristic of that sound, and two sounds which differ only in their relative degrees of consonance (or dissonance) are therefore functionally equivalent, or potentially so. This interpretation is consistent with our understanding of the meaning of dissonance in traditional harmonic practice, and with the fact that the music of Schoenberg and the other composers with whom we will be concerned here represents a more or less complete suspension of traditional harmonic procedures. The functional distinction between consonance and dissonance was one of the essential features of the tonal system of the 18th and 19th centuries, and one natural result of the suspension of that system would be the breakdown of this functional distinction.

The parallel between this equivalence of consonances and dissonances (as I interpret Schoenberg's statement) and my own principle of equivalence involves more than the idea of equivalence that is common to both. There is a further similarity in that Schoenberg's "consonances" are analogous to the simpler, aurally analyzable ("comprehensible") chords mentioned earlier, and his "dissonances" correspond to the more complex sound-elements, or the indefinitely-pitched noises. One of my first descriptions of the latter types of sound referred to "tone-clusters and other highly dense and dissonant chords," and indeed there is an obvious relationship—both acoustically and psychologically—between dissonance, complexity, and noise.

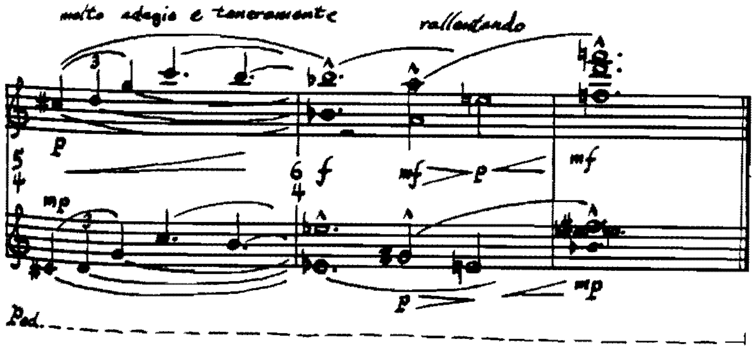
The kind of equivalence I am suggesting, however, is perhaps not a "functional" one in quite the same sense as is the equivalence of consonances and dissonances described by Schoenberg. It might rather be called a "substantial" or "material" equivalence, meaning not that these different kinds of sound necessarily have equivalent functions or musical effects, but simply that they have an *equal potentiality* for use as elemental building-materials in music. Thus the conceptual framework proposed here will not begin with *tones* as the primary units of the material—even though this might seem to be the logical starting point from an acoustical point of view. Rather it will postulate *sounds* and *sound-configurations* as its primary units, deriving this premise from psychological or more directly musical assumptions.

So far, we have been considering sound-elements of varying degrees of complexity in the vertical dimension, with no reference to their possible changes in time. But such sound-elements must also be examined in relation to the time-dimension, since they all have some extension in time, and their vertical characteristics usually vary with respect to time. This will lead to an expansion of the principle of equivalence to include sounds with considerable variation in time, and it will be seen that these, too, can function as basic elements in the larger sound-configurations or musical ideas.

But first, it should be noted that although no sound is time-independent in its acoustical features, we are not always aware of the changes that may actually take place in a sound. Even the simplest tone has a characteristic *time-envelope*, consisting of three different stages: an attack, a steady-state portion, and a decay in amplitude. But whether or not we actually perceive such changes is strongly determined by the musical context in which the sound occurs, and to some extent by conventions and listening habits. It is well known, for example, that the tone of the piano begins to decrease in amplitude almost immediately after the hammer strikes the string—piano-tone has, in fact, no steady-state stage at all—and yet we are virtually unaware of this when we listen to most piano music. This is strikingly demonstrated by reversing the direction of a recorded tape of piano music. The whole gestalt-character of the sound is altered quite drastically, and seems to bear not the slightest relation to the character of the original sound. During such an experiment one suddenly becomes intensely aware of the envelope of each tone, though it is merely the same envelope in reverse. In the case of piano-tone, it would seem that our awareness has been dulled by familiarity, but of course musical context has played its part here too. Most music for the piano has been written as though the tone did *not* fade away immediately, or it has been composed in such a way as to disguise this fact as much as possible. Playing techniques

have been conditioned by this fact too, as for example the technique of overlapping successive tones in a line in order to simulate a *legato* that is only really possible on instruments which can sustain a tone at a given dynamic level.

In some cases, however, the musical context does encourage an awareness of the envelope or variations in dynamic shape of the sounds, by the exploitation of the various possibilities of "touch" with the piano, for example, or of different kinds of articulation in other instruments. Such varieties of "touch" or articulation are—physically—nothing more than ways of varying the time-envelope of the sound. But, if they are perceived at all, it is usually as differences in the "quality" of the sound, rather than as dynamic variations *per se*. The time-envelope may become quite perceptible (whether apprehended as variations in loudness or as tone-quality) when the perceptual scale of the music is reduced in such a way as to encourage the perception of smaller details, as it is in much of Webern's music, and in certain pieces by John Cage (particularly those for "prepared piano"). But there are cases where even this reduction in scale in not necessary. In the example from the Ruggles piece (example 5) the listener is clearly intended to hear not only the fading away of the sound after the last chord has been struck, but also a kind of play of interference among several tones in the chord, whereby they seem to swell and fade and swell again, each at a different rate, so that now one is the loudest, now another, resulting in an effect of internal melodic movement. The sound is very much like that of a bell, whose inharmonic upper partials "beat" with one



Example 5. Carl Ruggles, "Evocation IV" (ms. 30-32).

another in a similar way, so that what one hears are changes in the pitch-structure of the sound with time, as well as the change in dynamic level.

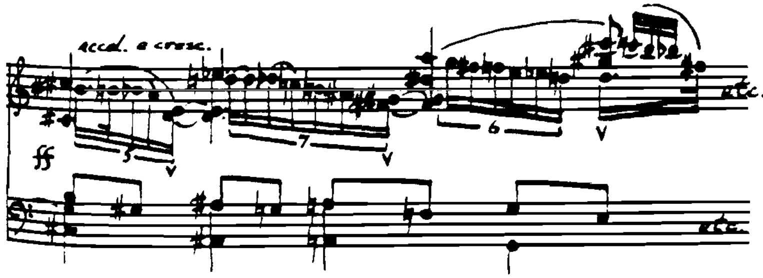
While the variations in amplitude mentioned previously were on the borderline between the realms of perceptibility and imperceptibility, the time-variations in the Ruggles example are clearly perceptible. And we can move gradually and by degrees into situations in which there can be no doubt that a sound's variations in time are no longer "subliminal," but in which the sound may still only have the character and function of a basic *element* in the larger configuration or sound-idea. Trills, tremolos, and fast repeated-notes fall into this category, as do certain kinds of arpeggiations, repeated figures, fast scale-passages and the like. (See examples 6 to 8.) They will have the "character and function" of basic elements when—because of the musical context—they are effectively "absorbed" into a larger configuration, or when their function within the configuration is made to be similar to that of their more static counterparts (i.e. trills and repeated-notes like sustained tones, tremolos and arpeggios like sustained chords, etc.). Now it must be said that these sounds which vary so with time are *not* identical to their "static counterparts," since there is always some reason (usually rhythmic) why one form of the sound, rather than another, is used in a particular passage—they are not interchangeable. But I suggest that they may be considered "materially equivalent," in the sense defined earlier, as having equal potentiality of serving as basic elements in the larger sound-configurations which constitute the musical ideas of a piece of music.

If we shift our attention now from the basic elements to the larger configurations themselves—configurations which would approximately correspond, in length, to the motives and phrases of an earlier music—it becomes apparent that the nature of such sound-ideas will be affected by the variety and complexity of the materials of which they are composed, as well as by the variety and complexity of arrangement or organization of these materials. Before examining such sound-ideas, it seems advisable to review some of the many factors which contribute to this "variety and complexity" in a more general way.

There are two factors that are particularly important in this respect: these are (1) the extension of the gamut or range of possibilities within nearly every one of the various parameters (i.e. pitch, loudness, timbre, temporal density, etc. [2]), and (2) a faster rate of

Example 6. Béla Bartók, Fourth String Quartet, I (p. 10).

Example 7. Arnold Schoenberg, op.11, #2 (p. 10).



Example 8. Charles Ives, "Concord" Sonata ("Emerson"), (p.17).

change in parametric values. [3] These two factors are related, in that a faster rate of change will generally mean the coverage of a greater range within a given time-span. With respect to certain parameters there has been both an extension of the range and an increase in the rate of change, while in others only the latter has taken place in any very significant way. The dynamic range, for example, can hardly be said to have been extended in any absolute sense—at least not since Beethoven, whose highest and lowest dynamic levels are comparable to those in 20th-century music. But there was surely never as high a rate of change of dynamic level as we find in the music of our time. The situation is similar for the time-dimension, too. Contrasts of temporal density have become a prominent feature of music, and again it is the increased rate of change in temporal density that is most noticeable, rather than the absolute range of differences between the slowest and the fastest extremes.

The asymmetrical phrase-structure which is so characteristic of 20th-century music can be viewed in this light, as also the more prose-like rhythmic development which it engenders. These are partially the result of the often noted tendency to avoid exact repetitions, and of a desire to replace the measured simplicity of verse and dance-rhythms with the freer rhythms of speech—and thus represent to some extent developments of rhythm for its own sake. But these asymmetries are also determined by the generally increased rate of change in other aspects of the music. That is, they are determined by the great variety, in both shape and substance, of the successive sound-elements and configurations in the music. There is often a continual change in the vertical density, for example (a two-part texture may be followed by one of six or eight parts; a narrow spacing may suddenly be replaced by a wide distribution of tones, etc.)—and this variety seems to necessitate a corresponding variety in length. It finally becomes difficult or even meaningless to speak of "phrase-structure" at all, and new terms will be needed for these sound-configurations that will make allowance for this greater variety in length, as well as in shape and "substance" or material.

Like loudness and density, pitch and timbre have also undergone a development in the direction of increased rate of change in parametric values. A characteristic feature of the melodic writing of many 20th-century composers—the use of wide skips or larger intervals at the expense of the smaller diatonic intervals—can be interpreted in this way. This, and the general tendency to employ the full range of a given instrument or voice, means covering more of

the pitch-compass in a shorter span of time—and thus an increased rate of change in the pitch-parameter.

But in addition, the absolute ranges of both pitch and timbre have been extended considerably. With regard to pitch, for example, it may be noted that the instruments sounding in the extreme high or low registers are now less often used merely to "double"—at a higher or lower octave—parts principally carried by the more standard instruments of the middle range of the pitch-compass. These previously "auxiliary" instruments have acquired a much greater independence within the total ensemble, and there is thus a widening of the effective "field" of pitch-events as such (as distinct from such elaborations of sonority as these doublings).

The use of the full range of an instrument—and, more specifically, the use of the extreme registers of an instrument—is also one of the ways in which the timbre-range has been extended. Other extensions include the employment of special techniques such as *sul ponticello* and *col legno* in the strings, fluttertongue in the winds, brass mutes, trombone glissandi, etc. as well as an increased use of the percussion battery of the orchestra.

An increased rate of change of timbre has also become a common feature of the music of our time, and the following statement by Schoenberg is instructive in this respect.

> "It is true that sound in my music changes with every turn of the idea—emotional, structural, or other. It is furthermore true that such changes occur in a more rapid succession than usual, and I admit that it is more difficult to perceive them simultaneously. . . . But it is not true that the other kind of sonority is foreign to my music." [4]

By "sound" he means what I am calling *timbre*—instrumental tone-quality—and the "other kind of sonority" would refer to a kind of musical texture in which the timbre does *not* change "with every turn of the idea." The comparison is with an earlier music and a more conventional instrumental style, and the question arises here whether the difference between the two kinds of "sonority" is simply a difference in degree, or one in kind. I think it is a difference in kind, and that the distinction he makes is fully justified. A 19th-century orchestral piece may show a great variety in timbre, and even perhaps a relatively fast rate of change in this respect, but the changes are seldom "with every turn of the idea"—which I take to mean *within a single idea*—but occur, instead, *with the appearance of each new idea*, in most cases. There is usually a high degree of timbral homogeneity within the limits of a single musical idea, and this is because the primary shaping-factor in these configurations is usually pitch, not timbre. If these represent, then, two different kinds of "sonority," it is nevertheless true that the development from the earlier one to the later one was a gradual process, moving by degrees, and that it would be difficult if not impossible to find any sharp line of division between the two stages of that process. But there are surely many natural processes which show a complete metamorphosis from one form to another, yet in which there is no perceptible break in the process itself, or in its evolution.

With this interpretation of Schoenberg's statement, we perhaps have a key to the solution of a problem that is raised by all these innovations which have been described here under the general categories of extensions of range and increased rate of change in the music. "I admit that it is more difficult to perceive them simultaneously," says Schoenberg, about the fast changes in "sound" in his music, and it might be said not only of timbre, but of all the other parameters of musical sound in which there has been this expansion of the range of possibilities—and not just about Schoenberg's music, but about 20th-century composers in general. One result of these innovations is the impression of discontinuity that the listener often receives on the first hearings of a piece, and an important question is raised: how or where is one to find that thread of continuity which we assume to inhere in every integral work of art? I think the answer to this question involves the ways in which the ear and mind organize the component sound-elements into larger units or gestalten, and this will depend upon both the way one listens, and upon the actual configurations in the music.

The last problem of the "actual configurations" will be studied in more detail in Section II, but here a few things might be said about "the way one listens." It seems to me that the first step in the direction of finding continuity amid the apparent discontinuity produced by these extensions of range is the acceptance of the wider gamuts as in some way normal, admitting the new events occurring in the extreme "registers" of each parameter to be within the "range of possibilities," rather than outside of it. This may seem to involve nothing beyond the assimilation of the "novelties

of a new style" mentioned at the beginning of the book, but it is more than that, and is a factor that must be considered in our attempts to arrive at a meaningful basis for musical description and analysis. The second step involves an understanding of the relative nature of continuity and discontinuity, and of some of the factors causing this relativity.

The relativity of continuity and discontinuity might best be illustrated by an analogy with a similar situation in the realm of vision. It often happens that one's first impressions of a modern painting do not correspond with one's later impressions, or with the intentions of the painter. At first one may see an apparently random distribution of colors, shapes or lines, only later discovering a figure perhaps, or objects of a still-life, or elements of a landscape. At some point in the process of studying the painting the seemingly random elements are subjectively integrated, making perceptible the configurations that are essential to one's understanding of the work. In the terms of the previous discussion, we can say that a continuity has been found within what at first seemed a condition of discontinuity; relations are perceived among elements that had seemed disconnected and unrelated.

Now what are the factors leading to the discovery of continuity—factors whose negative effect is to prevent this discovery? One such factor has already been discussed—the "mental set" which can cause events occurring in the extreme ranges of each parameter to interrupt the sense of continuity. But there are two other factors which are even more important than this one, and these are the factor of *scale*, and that of *focus*. There are at least two forms of the latter, and I will consider these first, before examining the question of scale. The two forms are (1) *textural focus*, and (2) *parametric focus*. The first is the most obvious, and little need be said about it, except that if one's attention is directed towards one or more of the *less* essential parts in a complex texture, the *more* important structural features of the larger configurations may be missed. This assumes, of course, a situation in which there *is* a hierarchy of more and less essential elements—which may not always be so—but the situation does occur often enough to make this a factor worth considering. In the final analysis, perhaps, the very richness of a work of art—in any medium—may be due to the ambiguities it allows in this respect, and to the possibility of directing the attention toward the secondary elements and finding these meaningful. But in the beginning, at least, there must be some reckoning of what the most important parts might be.

Parametric focus is analogous to textural focus in many ways, but it is something different, and perhaps not so obvious as the latter. In the course of this book, an attempt is made to demonstrate the greater importance that has been given in 20th-century music to *all* the parameters of musical sound; that whereas in earlier music the responsibility for the articulation of musical ideas was mainly given to the pitch-parameter, the other parameters have begun to carry more and more of this responsibility, sometimes even to the extent of replacing the function of pitch altogether. It is further suggested that the relative *degree of articulation* in the several parameters (one manifestation of the rate of change discussed earlier) may vary—and with that, the parametric focus will vary—not only from one piece to another, but within the same piece, or even within a single passage in a piece. If this is so, "the way one listens" to the music is certainly going to be affected. Such changes in parametric focus will require a corresponding flexibility on the part of the listener, and it will be necessary to acknowledge the possibility of these changes of parametric focus or parametric articulation, and to allow for them in our conceptual approach to the music. It is partially the failure to do this which has led to the attitude so often encountered in criticisms of some 20th-century techniques, which would reduce them to "mere color-effects," or "purely rhythmic experiments," etc. The listener who can accept only pitch as a primary shaping factor in the articulation of musical ideas is bound to hear "empty spaces" in much of the music of the 20th century, and may eventually have to reject altogether some of the more advanced expressions of the musical art, such as Varèse's "Ionisation," for percussion instruments, the pieces for "prepared piano" of John Cage, electronic and tape-music, etc. This is unfortunate and unnecessary, when all that is required to include such music within the larger "main stream" of musical development is a broadening of our conceptual framework so as to include such phenomena as this change of parametric focus.

That factor in the creation of apparent discontinuity which I have called *scale* is even more important than either textural or parametric focus, and will lead us more directly to the essential point of this section of the book. I am not using the word *scale* in the ordinary musical sense here, but rather in the sense a draftsman or map-maker might use the word, and more generally, as it is used in

the visual realm, from which the best illustration may again be taken. We know from our visual experience that a change in scale of a picture of a thing, or a change in the distance from which we view a thing—whether it be a picture, a landscape, or the figure of a person—can substantially alter the total impression we will have of it. The overall gestalt-character of the thing seen is thus to a great extent determined or conditioned by the scale on which we view it, and this depends not only on physical conditions such as size and distance, but also on the mental set and purposive attitudes of the viewer. If we imagine again the situation described before—a person whose impressions of a painting are still disconnected and unrelated—it is apparent that the configurations he does perceive may be only the *details* of a larger configuration, and that his attention to these smaller units may actually prevent his perception of the larger and more essential configuration. The process also works in the reverse direction—the larger units being mistaken for detail—in which case the whole structure must inevitably seem incomplete. The full range of this process might be illustrated by imagining a scene—say a field of wheat—which from a certain distance will appear continuous, having a homogeneous texture that is unbroken by contrasting elements. If one moves closer, this texture will gradually become less and less homogeneous, until at last the distance is so shortened that one's field of vision can only encompass a few of the elements—the stalks of wheat. At this point, those elements which before had been absorbed into the larger unit—perceived as texture, but not distinguishable separately—become whole units in their own right, and the spaces between them are seen as real breaks in continuity. Similarly, if one starts from the original vantage-point and increases the distance from the field, one will eventually reach a point where the whole field is only an element in a larger scene—a larger gestalt—that includes houses and a road perhaps, and other fields of a different color or texture. Again a continuity has been replaced by a relative discontinuity.

If we transfer this now to the realm of musical perception, it should be evident how it applies to the problem of apparent discontinuity in music, and of the relativity of continuity and discontinuity. If the scale on which the listener is prepared to grasp successive sound-configurations is not commensurate with the scale on which the music is actually organized, there will be a greater sense of discontinuity than is actually implicit in the music. If the music is highly complex, with many and variegated elements contained within the limits of each musical idea, such a listener will be in the position of the viewer described above whose attention is fixed on the details, being thereby unable to see the larger configurations of the picture. Or he will be like a person learning a new language, who misses the sense of a sentence heard in that language because his mind has "stopped" to translate the first or second word of the sentence. Here again, an undue attention to the elements has prevented the apprehension of the larger configuration as a singular gestalt. This kind of situation is most likely to arise in music like that of Schoenberg or Ives, which usually requires the simultaneous perception of far more elements than does the music of most other composers. But in general, 20th-century music is far more demanding in this way than earlier music was.

In much of the music of Webern, however, we find just the reverse situation. Here there is a very different scale of musical organization, demanding a different scale of perception, in that small sound-structures—which in most other music would be no more than elements that are not intended to be heard separately—become with Webern the essential musical ideas, primary musical gestalten that must be perceived as relatively complete or self-sufficient in themselves. Here the result of a disparity between the scales of the composer (i.e. of the music) and of the listener will be a sense of incompleteness, if not of discontinuity.

Finally, and no less important than the above, it should be noted that the scale of organization of the successive musical configurations in any single piece of music may change considerably from one to the next, and this requires a greater flexibility of the listener's scale of perception. The difference between 20th-century music and earlier music, with respect to this variability of scale, is similar to the difference between the two kinds of "sonority" described earlier. The development has been a gradual one, but it becomes a thing of a different kind in the music of this later period. In 18th- and 19th-century music such variations could generally be referred to some approximate standard or norm, and in fact, the important structural potentialities of such variations owe their strength to the very existence of such a norm. These norms no longer function in contemporary music, however, and the range of variation is much greater, so that *variability itself* must be recognized as a kind of norm. This last statement obviously applies not only to variability of scale, but to the other innovations discussed so far as well—change of textural and parametric focus,

the faster rate of change of parametric values, and the extension of the ranges in the various parameters. To a great extent, the impression of discontinuity and other "sense-interrupting effects" may be reduced or neutralized by the mere acceptance of such variability as normal. And, as it is with perception, so it must be with analysis and description, and a conceptual framework is needed which will allow for all these new possibilites. Only with such a broad conceptual framework as a basis, can we proceed to an analysis of the specific structural forces which are active in 20th-century music.

The recognition of the variability of scale with respect to the larger sound-configurations or musical ideas leads to a final extension of the principle of equivalence to make it applicable now not only to the component elements of sound-configurations, but to these larger configurations themselves. That is, we must admit a "material equivalence"—with respect to their potential function (as musical ideas)—of a much greater variety of sounds and sound-configurations than would have been justified or necessary in pre-20th-century music. I say "sounds and sound-configurations" here advisedly, because—as was pointed out about the reduced scale of organization in the music of Webern—relatively simple sounds, which in another music might be only elements, are sometimes capable of functioning as musical ideas in their own right. Recalling now what has already been said about the greater range of complexity of sound-elements, it should be apparent that there is some degree of overlapping between the range of elements and the range of sound-ideas, and the principle of equivalence must now be understood to include this ambivalent potentiality of sounds and sound-configurations which fall within the overlapping portions of their respective ranges.

Whether a given sound or sound-configuration is to be considered merely as an element or as a more self-sufficient musical idea depends almost entirely upon the musical context in which it is heard. There is virtually no objective characteristic of the sound itself (except duration) which can show the analyst in which of these two categories it ought to be placed. Only its function within the larger design can reveal this—its relation to other sounds and sound-configurations. But the study of such relations, and thus the study of function, cannot begin without some definition of the *things* involved in the relations—the entities that are functioning—the sounds and sound-configurations themselves.

As a result of this last extension of the principle of equivalence, the distinction between element and "idea" has been relegated to the realm of context. The distinction thus qualified, the question arises as to what characteristics are held in common by all these sounds and sound-configurations which have been the subject of our analysis so far. It will be seen in the course of this book that there are many specific features which may be involved in an answer to this question, but the most general characteristic common to them all—one which has always been at least implicit in the previous discussions—is the fact that they are *perceived as units*. Almost by definition, the sounds and sound-configurations we have been dealing with here exhibit that unity or singularity which—in the visual domain—is characterized by the term "gestalt," and it is evident that some consideration ought to be given to the principles of gestalt perceptual psychology, in our search for an expanded conceptual framework for 20th-century music. In his *Principles of Gestalt Psychology* [5] Kurt Koffka says:

> "The laws of organization which we have found operative explain why our behavioral environment is orderly in spite of the bewildering spatial and temporal complexity of stimulation. Units are being formed and maintained in segregation and relative insulation from other units. . .without our principles of organization. . .the phenomenal changes produced by these changes of stimulation would be as disorderly as the changes of stimulation themselves. . .order is a consequence of organization, and organization the result of natural forces."

This statement has an obvious relevance to the musical problems we have been considering here, and in the next section of the book I shall try to demonstrate the applicability of some of these same "laws of organization" to musical perception. At this point, however, I want to emphasize that the first condition mentioned by Koffka for the appearance of order, within a "bewildering complexity of stimulation," is the perceptual formation of *units*, "maintained in segregation and relative insulation from other units." This will be a basic assumption in all the arguments that follow. And one of the first questions which must be asked about the various sounds and sound-configurations that occur in music is: what factors are responsible for their unity or singularity, and what factors effect their "relative insulation from other units?"

To facilitate the examination of such questions, I shall introduce here a few basic definitions—or rather, some new terms which may serve as points of departure for further definitions and distinctions. The continued use of such terms as "sound-configuration," "musical idea," etc., seems to me unsatisfactory, the former being too general, the latter too specific, and it would be misleading to try to adapt familiar terminology to the purposes of this investigation. Words like "phrase," "theme," "chord" and "chord-progression," and even "melody" and "harmony" would have to be so reinterpreted that they would cease to have much meaning. I have rather attempted to develop a terminology which would be specific enough to make significant distinctions possible, and yet remain general enough to allow for some degree of inner expansion.

In place of "sound," "sound-configuration" or "musical idea" (as these have been used up to this point in this book), I propose the word *clang*—to be understood to refer to any sound or sound-configuration which is perceived as a primary musical unit—a singular aural gestalt. For the subordinate parts of a clang, I shall continue to use the word *element*—whether these are articulated in the vertical dimension as "linear" or *concurrent* parts, or in the time-dimension as *successive* parts—i.e. tones, chords, or sounds of any kind. Finally, some term is needed to designate a succession of clangs which is set apart from other successions in some way, so that it has some degree of unity and singularity, thus constituting a musical gestalt on a larger perceptual level or temporal scale—though it will not be as "strong" a gestalt (a term used by Köhler [6]) as is the clang. For this larger unit I shall use the word *sequence*, and further distinctions as to type and function will be made after an examination of its most general characteristics, in Sections II and III of this book.

I have adopted this word *clang* for several reasons, and some explanation of these reasons may help to clarify for the reader my understanding of the term. First, its only current meaning in English ("a loud ringing sound, as of metallic objects struck together"—Webster) suggests a kind of sound or sonority—complex and dissonant—which is frequently to be heard in 20th-century music, and the consideration of which first led me to the reexamination of musical materials and formal factors as outlined in this book. Second, although the word has had some currency in English

(British, not American) writings on acoustics, it does not seem to have been used very widely or over a very long period of time with any single meaning. It is sometimes used in such writings on acoustics to mean a compound tone (i.e. one composed of several harmonic partials), but at other times it is used to mean the sound of an interval or chord. My definition of the word might be considered an extension of these meanings to include *any* singular sound or sound-configuration. Third, its derivation from or association with the German word, *Klang*—meaning both "sound" and "tone"—carries with it some implication of the notion of equivalence described earlier. And finally, *clang* is a word that refers specifically to *auditory* perception, and not—like so many others that we use or may be tempted to use (such as "configuration," "pattern," "object," "idea," etc.)—borrowed from the visual or other perceptual realms.

The distinction between *clang* and *sequence* is intended primarily to be a generalized functional distinction, and will not always be entirely clear-cut or unambiguous, in actual musical examples. But in general, the clang is a sound or sound-configuration that is more or less immediately perceptible as an aural gestalt, while the sequence—being apprehended in a less immediate way than the clang—would be what Köhler called a "weak gestalt." Similarly, the distinction between an element and a more complete or self-sufficient clang will always be a relative matter—the element being, in a sense, a "smaller" clang that is effectively absorbed into a larger clang, thereby losing much of its individuality as a musical gestalt.

It should be evident then, that although the clang may often correspond in length or character to the motives or the phrases of traditional music, the word is not meant to define a structural or formal type at this perceptual level, as do the words "motive" and "phrase," but rather a kind of musical event and perceptual situation that may involve many other types of sound-structure than these. The only thing that is common to them all is their perceptual immediacy, and their singularity—i.e. their character as aural or musical gestalten. The principle of equivalence may now be understood to mean that virtually any sound or sound-configuration—no matter how simple or complex it may be from an acoustic point of view—may function within the larger musical context as a clang, if only it is *perceived* in that context as a primary musical gestalt.

There are some important similarities between this concept of the clang or aural gestalt and Pierre Schaeffer's *"objet sonore"* (or,

more specifically, the kind of sound-object he calls the *cellule"*—and I must acknowledge here my indebtedness to the writings of Schaeffer, [7] in the initial development of the ideas presented in this book. The *objet sonore* is defined as practically any sound or series of sounds recorded on disc or tape (within certain obvious limits of duration, of course), so that the compositional process automatically involves the "potential equivalence" of various elements, as this has been described here, as well as certain implications of gestalt-character with respect to the sounds.

But there are also some significant differences between Schaeffer's ideas and my own, and these should be noted here along with the similarities. Schaeffer's definitions are generally "operational" definitions, to an extent that tends to restrict their applicability to the particular medium with which he is working—"*musique concrète*," the compositional organization of recorded sounds on tape. The techniques of "transmutation and transformation" which he employs clearly involve the possibility that the same "sound-object" may function at one place in a composition as a clang, at another as an element, or even as a sequence, and it may be split up or rearranged in ways that completely alter its original gestalt-characteristics. Thus Schaeffer's definitions refer less to the perceptual events in the music (or rather, in the musical experience) than to the physical or acoustic materials that are manipulated in the process of composition. And it is for this reason, perhaps, that he has emphasized the differences between the "abstract" music of the past—including even most 20th-century music—and his own *musique concrète*. I think the essential difference between them is not a musical difference, however, but a *technical* one, and—from the purely musical standpoint—hardly justifies such a distinction in name, as between "abstract" and "concrete."

From a broader point of view, it has always seemed to me that the major innovations in 20th-century music have tended from the very beginning to involve something like the "sound-object"—if this is interpreted as an "object" of perception, rather than an object of technical manipulation. The concept of the clang, therefore, might be considered an outgrowth of Schaeffer's *"objet sonore"*—but directed toward the perceptual event itself, rather than the acoustic source of that event. Thus, the clang-concept should be applicable to music in any medium, whether instrumental or electronic, whether it employs natural or synthetic sounds, whether its psychological implications are "abstract" or "concrete."

Beginning then with the definitions of *element*, *clang* and *sequence*, and particularly the definition of the clang as *a sound or sound-configuration which is perceived as a primary musical unit or aural gestalt*, I shall try in the next section of the book to answer the following questions: (1) what factors are responsible for the unity or singularity of the clang?—and the necessary corollary to this—(2) through what factors is one clang "segregated" from another in the sequence?

Section II.

"THE TWO-OR-MORE-DIMENSIONAL SPACE IN WHICH MUSICAL IDEAS ARE PRESENTED IS A UNIT. . . . All that happens at any point of this musical space has more than a local effect. It functions not only in its own plane, but also in all other directions and planes, and is not without influence even at remote points. . . . The elements of a musical idea are partly incorporated in the horizontal plane as successive sounds, and partly in the vertical plane as simultaneous sounds. . . . Every musical configuration. . .has to be comprehended primarily as a mutual relation of sounds, of oscillatory vibrations, appearing at different places and times."

Arnold Schoenberg, *Style and Idea*, pp. 220-223.

"The first phase of apprehension is a bounding line drawn about the object to be apprehended. An esthetic image is presented to us either in space or in time. . . But temporal or spatial, the esthetic image is first luminously apprehended as selfbounded and selfcontained upon the immeasurable background of space or time which is not it. You apprehend it as *one* thing. You see it as one whole. You apprehend its wholeness."

James Joyce, *A Portrait of the Artist as a Young Man*, p. 212.

"The form, then, of any portion of matter. . . and the changes of form which are apparent in its movements and in its growth, may in all cases alike be described as due to the action of force. In short, the form of an object is a 'diagram of forces,' in this sense, at least, that from it we can judge of or deduce the forces that are acting or have acted upon it. . ."

D'Arcy Wentworth Thompson, *Growth and Form*, p. 16.

Section II.
Gestalt-Factors of Cohesion and Segregation.

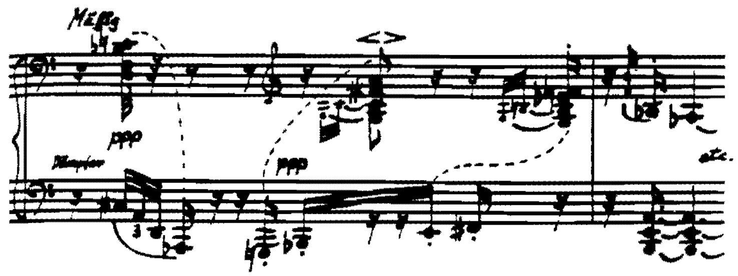
In 1923, Max Wertheimer published a paper entitled "Laws of Organization in Perceptual Form," [8] in which he demonstrated certain factors of unit-formation and segregation, operating within systems of points and lines in the visual field. This paper has since become one of the cornerstones of gestalt psychology. Wertheimer's procedure was simple, but none the less elegant in the way each of the various cohesive factors was isolated from the others, and shown to be capable of functioning independently. In the course of the demonstration, frequent analogies are suggested to auditory configurations, but no attempt was made to analyze this realm of perception in any thoroughgoing way. And in general, the gestalt psychologists' studies of perception have been directed primarily to visual problems, probably owing to the greater directness and immediacy with which visual forms may be presented, perceived and described. Nevertheless, many of the principles of organization of visual forms may be shown to be involved in auditory perception, often with no more than a simple translation of terms. In other cases, the problems are not so simple, but the writings of the gestalt psychologists, [9] Wertheimer, Koffka and Köhler in particular, can still serve us as a guide and precedent.

The first factor demonstrated by Wertheimer was called the *factor of proximity*, and might be stated as follows: in a collection of similar visual elements, those which are close together in space will naturally or spontaneously tend to form groups in perception—other factors being equal. A very simple example showing the effect of relative proximity on visual grouping is shown in Figure 1, below.

00 000 00 0000 00 000

Figure 1

The analogy in musical perception is obvious, when we substitute time for space, and sound-elements for visual elements—in the statement given above. In example 9, for instance, the sounds which are separated by the shortest intervals of time (including those



Example 9. Arnold Schoenberg, op.11, #3 (m. 22).

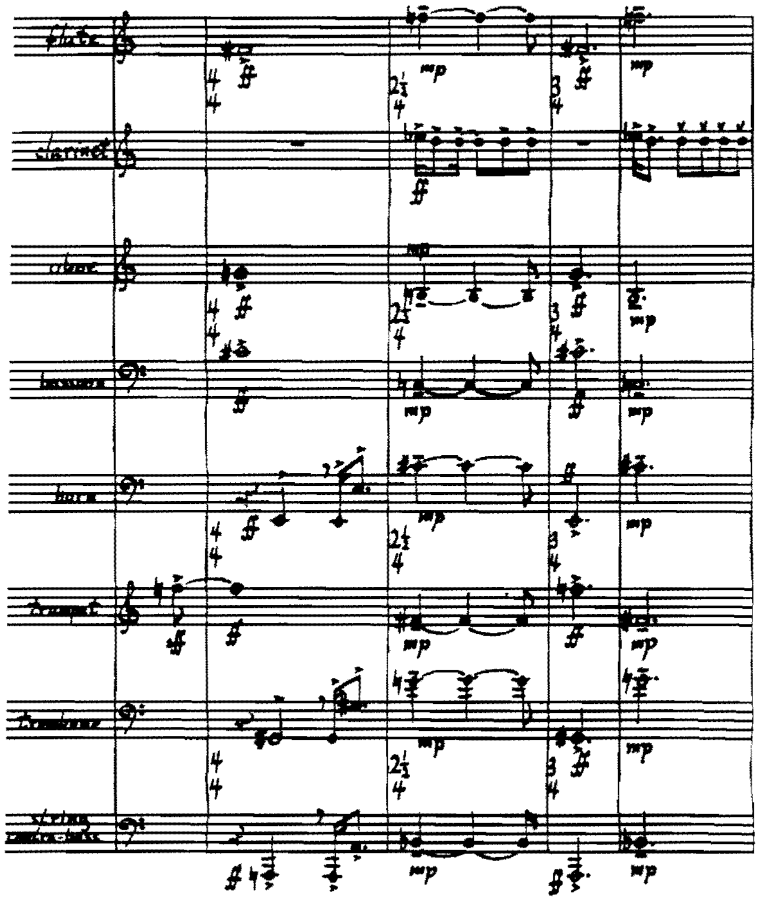
sounding together, of course) tend to form units or groups, while the longer time-intervals (in this case, the silences) cause unit-segregation. It can be seen from this example that temporal proximity may be manifested in either (or both) of two ways—as *contiguity* or as *simultaneity*. The essential principle is the same in either case. Applied to auditory or musical perception, the *factor of proximity* might be formulated as follows: *in a collection of sound-elements, those which are simultaneous or contiguous will tend to form clangs, while relatively greater separations in time will produce segregations*—other factors being equal. (The "other factors being equal" clause is very important, as will soon become apparent.)

A second factor in the formation of visual groups Wertheimer designated as the *factor of similarity*. In a collection of visual elements, those which are similar will tend to be grouped by the eye, as is shown in Figure 2, below, in which the elements are equally spaced so that the proximity-factor can have no effect on the grouping.

0 0 0 # # # 0 0 0 0 # #

Figure 2

The same principle in musical perception relates to the fact—well understood by any musician, at least implicitly—that sounds played on the same instrument (i.e. of similar timbre) or in the same pitch-register (of similar pitch) tend to seem "connected" and to form groups more easily than sounds that are relatively dissimilar in these respects. Examples 10 and 11 represent two typical configurations in which relative similarity of pitch (ex. 10) and of timbre (ex. 11) is the primary determinant of coherence within each clang. In the Varèse example, the pitch-similarity between the F in the trumpet and the Eb-D in the clarinet is such a strong cohesive factor in this linear element of the larger clang that it overcomes the segregative influence of timbre-difference between the two instruments. Thus, one does not hear as a unitary element the F-F#-F-F#. . .being played by the trumpet, but rather a single line that passes from trumpet to clarinet. In the Webern example, on the



Example 10. Edgard Varèse, "Octandre", II (mm. 50-53).

Example 11. Anton Webern, op.10, #2 (beginning).

other hand, the effect of pitch-similarity is much less powerful than the timbre-similarity which unifies each of the two instrumental lines (i.e. Eb clarinet and violin) into singular units—and the difference in timbre which keeps them separate and distinct from each other, even though the parts cross melodically. And it is the change in timbre—from clarinet to oboe—that will effect the perceptual separations between clangs 1 and 4, in spite of the pitch-similarity between the end of the clarinet-line and the beginning of the oboe part. Thus, one parameter may run counter to another with respect to the operation of this factor of similarity. But it is the existence of a relatively higher degree of similarity in some parameter that is the unifying force in such clangs.

Note also that the cohesive force of the similarity-factor implies—as its necessary corollary—the segregating effect of dissimilarity, just as, with the factor of proximity, a greater separation in time (i.e. relative "nonproximity") will tend to cause segregation. The very process of unit-formation necessarily implies relative separation from other units—or from other parts of the perceptual field—and this fact will become more and more significant when we begin to analyze the possibilities for gestalt-formations on various perceptual levels or temporal scales.

The factor of similarity applies not only to pitch and timbre, but also to the other parameters—dynamic level, envelope, temporal and vertical density, etc.—and in fact it may be said to function with respect to any attribute of sound by which we are able, at a given moment, or within a given time-span, to distinguish one sound or sound-configuration from another. Thus, for example, morphological similarity or similarity of form among the component clangs of a sequence constitutes a powerful factor in the unification of that sequence.

Finally, it should be noted that the cohesive and segregative forces of relative similarity and dissimilarity apply not only to successive groupings—where, for example, one clang is segregated from the next clang in a sequence—but also to concurrent configurations, in which one clang is distinguished from another that is sounding at the same time (as was the case in example 11). The effects of the similarity-factor may thus run counter to those of the proximity-factor—and indeed, true polyphony would be impossible if the only conditions leading to clang-formation and segregation were contiguity and simultaneity.

We may now formulate the *factor of similarity*, with specific reference to musical perception, as follows: *in a collection of sound-elements (or clangs), those which are similar* (with respect to values in some parameter) *will tend to form clangs (or sequences), while relative dissimilarity will produce segregation*—other factors being equal.

Thus far in the analysis of these factors of cohesion and segregation it has been necessary to isolate each of them, and consider its effects separately. This is an abstraction, of course, and it should not be forgotten that in every real musical configuration both of these factors (and others, to be described in a moment) are operating simultaneously, although they do not usually exert equal force in any given configuration. In addition, they may be more or less co-operative, their results in perceptual organization varying over a wide range from complete congruency or mutual reinforcement, through partially ambiguous, overlapping effects, to completely ambivalent, multi-structural configurations produced by antithetical relationships between the two factors. In the latter case,

(which is not by any means an exceptional one, even in pre-20th-century music) a given collection of elements may be perceived in two or more different and distinct configurations, yielding, that is, two or more clangs simultaneously, each of which may be equally important in the larger musical context.

Although the factors of proximity and similarity are not the only ones involved in the organization of perceptual units, they are the most basic—i.e. the most effective—and the most frequently decisive in the determination of clang- and sequence-unity. For this reason, I shall refer to them as the *primary* factors of cohesion and segregation. In addition to these, there are four *secondary* factors which will be considered here. These are (1) the factor of *intensity*, (2) the *repetition-factor*, (3) *objective set*, and (4) *subjective set*. Before describing these, however, I want to introduce a very simple graphic representation, which can help to illustrate the factors of proximity and similarity, and perhaps also to clarify some points in the arguments that follow.

As shown in Figure 3, the horizontal axis of the graph represents time, and the vertical axis represents an ordinal scale [10] of values in one of the various parameters—i.e. in *any* parameter—it does not matter here which parameter is involved. If one plots, on such a graph, the variations in some parameter with time, the result will be what I shall call a *parametric profile* of the element, clang or sequence involved, which gives a general picture of the configuration with respect to that particular parameter. For example, if the vertical ordinate is pitch, such a plot will show melodic contour (but note that with the present definition of the vertical scale, the plot cannot tell one anything about the actual pitches or intervals in the configuration). If the vertical axis is made to represent loudness, one might plot the time-envelope of the attack and decay of a simple element, or the dynamic shape of some larger clang or sequence. Thus, such a graphic representation might be considered a kind of two-dimensional perceptual "model"—albeit a very primitive one—which can be used to depict one aspect of the perception of a given configuration—that aspect which corresponds to the variations in time of one parameter.

It will be evident that distances between individual elements in such a graph—when measured along the horizontal axis (or, more precisely, distances between their respective projections onto the horizontal axis)—will show their relative proximity in time. Similarly, distances measured in the vertical direction will indicate, in a general way at least, relative similarity or dissimilarity between

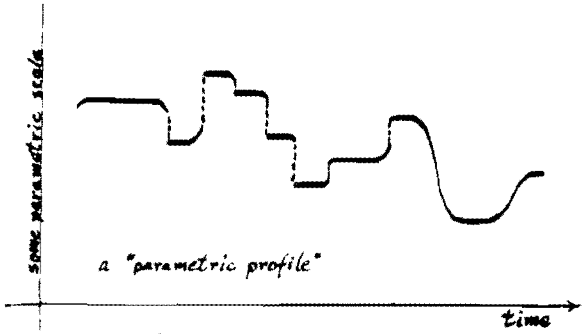


Figure 3.

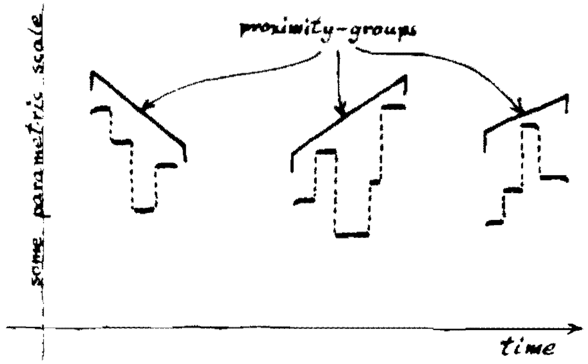


Figure 4.

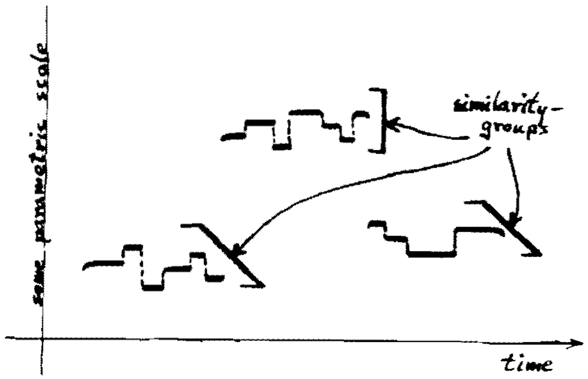


Figure 5.

these elements, with respect to the parameter designated in the graph. Thus, proximity in time is represented by proximity in space—measured horizontally—while parametric similarity is represented by proximity (in a sort of one-dimensional "attribute-space")—measured vertically. In Figures 4 and 5, two hypothetical configurations are plotted, the vertical axis being left unspecified as to the particular parameter intended, merely representing (as in Figure 3) any distinctive attribute of sound, in terms of which such an ordinal scale might be constructed. The configuration in Figure 4 would correspond to a situation in which proximity is the principal factor in the formation of groups, whereas Figure 5 shows unit-formations primarily determined by the factor of similarity.

The inherent two-dimensionality of such graphs imposes certain limitations on this "perceptual model," since the perceived form of every real musical configuration will involve an interaction of all parameters—not just one—and these parameters may not always be perceived independently, as this method of analysis of single parametric profiles might seem to imply. But by isolating the various parameters in this way, and considering each profile separately, it becomes possible to formulate certain general principles that will still be valid in more complex conditions that result from the simultaneous influences of several parameters in a clang or sequence.

The first of the secondary factors of cohesion and segregation—the factor of intensity—relates to the singular directionality of the parametric scales employed in the graphs. That is, we generally assume an absolute "up" and "down" on these scales—a higher and lower parametric value—which is somehow related to what might be called *musical* or *subjective intensity*. I say "somehow related" because, although this "directionality" is understood and utilized by the musician in practice—and is implicit in most of the devices employed by both the composer and the performer in creating climaxes, building up musical tensions, intensifying or activating a passage of music, etc.—I know of no attempt to define these conditions explicitly, much less to explain them in non-musical terms. It is a common fact of musical experience that a greater subjective intensity is usually associated with a rise in pitch, an increase in dynamic level or in tempo, etc. Similarly, a change from a "smooth" or "mellow" timbre to a "harsh" or "piercing" timbre, or from a more consonant to a more dissonant interval, is felt as an increase in subjective intensity.

An explanation of these conditions might eventually be derived from certain concepts of *information theory*, beginning with measures of the information transmitted in the form of neural discharges in the "communication channel" between the ear and the brain. Such measures have been made, at least for frequency and amplitude, and these indicate that a higher rate of transmission of neural information is indeed associated with both a higher pitch and a greater loudness, and some inferences from this data might be made in regard to timbre, vertical density, and perhaps other parameters as well.

But this can be no more than a beginning of an explanation, because many more strictly psychological factors may be involved, and if we had to wait for conclusive evidence in the form of physiological data we should probably never be in a position to describe this factor of subjective intensity in a satisfactory way. I shall, therefore, simply define an upward displacement on a parametric scale as a change in value in that parameter which produces, or is associated with an increase in the subjective intensity of the sensation. In addition, I shall call the measure of relative height on such a scale *parametric intensity*. Parametric intensity is thus to be understood as an approximate measure—in one "dimension"—of the more inclusive musical or subjective intensity of a perceived sound.

Consider then what happens when listening to a moderately complex clang. It may be observed that one's attention is not usually distributed evenly among the component elements, but is focussed more sharply on certain elements than on others. For example, in a clang with several concurrent elements—delineated, let us say, by separate instrumental parts—the attention is likely to be directed to that element which is loudest, or (if they are all equally loud) to the one with the most intense timbre, or (supposing all elements to be equal in both loudness and timbre) to the one that is highest in pitch, etc. In each case the attention will tend to be directed toward—and more sharply focussed upon—the element which exhibits the highest values on some parametric scale. If the difference in parametric intensity between one such element and the others is not too great, the result will be a variation in focal "resolve," with the most intense element being heard more clearly, seeming more immediately "present" in perception, while the less intense elements will be more or less "blurred"—more or less "remote" as perceptual objects. In this situation, I am assuming that all the elements are heard as parts of a single clang, in spite of the

dissimilarities between them, but of course, if there is too great a difference in parametric intensity between one such element and others, a subdivision may occur—as a result of our second factor of cohesion and segregation, the factor of similarity—so that one will hear two separate clangs instead of one.

So far, we have found nothing new in the way of grouping tendencies, but if the analysis of the intensity-factor is transferred now from the vertical to the horizontal dimension it will be found that this factor by itself can produce unit-formations in time—independently of the factor of proximity, and in a way that is not accounted for by the similarity-factor—as this has been formulated—although parametric intensity is obviously related to the question of similarity and dissimilarity of parametric values. I am referring here to what we call *accent*, and more specifically, to the *group-initiating tendency* associated with the accent. I suggest that similar conditions hold for the effects of intensity-differentiations *in time* as were observed above in the case of vertical differentiations, and that the same terms might be used to describe the perceptual results, if not to explain them. That is, in a succession of sound-elements showing marked variations in intensity (in some parameter), the attention will be more sharply caught by the more intense, accented elements, while the less intense elements will be relatively blurred, and—by way of memory, or perhaps through some kind of kinesthetic response-process—the attention at certain moments may actually be directed *backwards* in time, toward the most recent accented element, until a fresh accentuation redirects the attention into the more immediate, present moment.

Such a process might be illustrated graphically as in Figure 6, where each arrow represents a kind of "attention vector," associated with each successive element in the graph. The length of such a vector would indicate relative clarity, focal resolve, etc., while the direction of the vector would represent the direction or displacement in time of the perceptual attention at each occurrence of a new element. I have placed the origin of each vector at a point on the time-axis corresponding to the beginning of each new element. If one now drops vertical projections from the *upper* terminals of each vector, marking off the points of intersection of these projections with a third horizontal axis, the groupings resulting from the factor of intensity alone are again shown by the relative proximity of the points in space (measured horizontally)—just as actual proximity in time would be. Whether or not this corresponds to some kind of

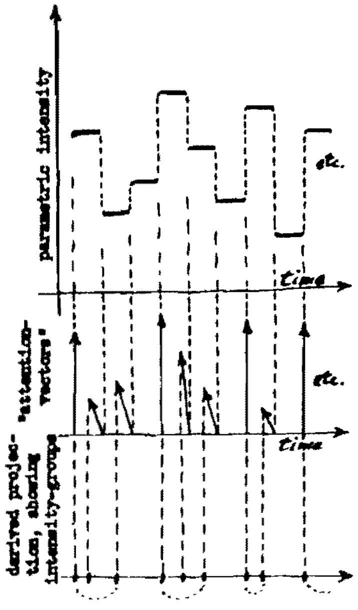


Figure 6.

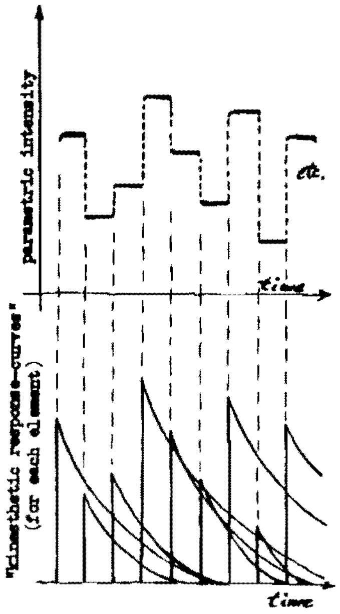


Figure 7.

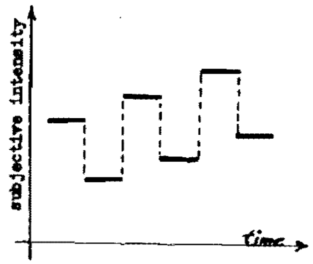


Figure 8.

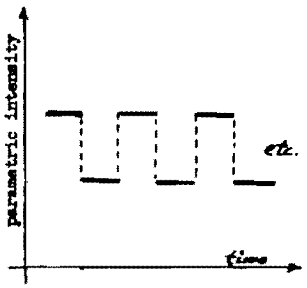


Figure 9.

distortion or "clustering" of successive moments in subjectively experienced time I have no way of knowing—and such an interpretation is not really necessary to the argument, although it does represent an intriguing possibility.

Although the above description of the grouping tendency of the intensity-factor has several advantages, it is not altogether satisfactory because of the speculative character of the subjective process

represented by the "vectors." Consequently, I shall offer two alternative hypotheses—equally speculative—which might account for the group-initiating effect of accentuation, either singly or in combination. The first relates the intensity-factor to the factor of proximity—interpreting it, in fact, as a special case of simultaneity—while the second would represent a special manifestation of the similarity-factor.

(1) The first hypothesis is based on the assumption that sounds evoke kinesthetic responses in the listener, the relative *durations* of which are in some way directly proportional to the parametric intensity of these sounds—the response to a more intense sound thus lasting longer than the response to a less intense sound. This may be represented graphically by means of a plot of the subjective intensity (or the magnitude of the kinesthetic response) versus time—arranged, as before (in Figure 6), "in parallel" with the plot of parametric intensity versus time. This is shown in Figure 7, (using the same parametric profile as in Fig. 6), and it will be seen that the appropriate unit-formations are indicated in the lower plot by the way in which the response-curves for the more intense elements tend to overlap and "absorb" those for less intense elements. The perceptual result of such a situation would be a degree of subjective "simultaneity" which would tend to favor groupings initiated by the accented elements.

(2) The second alternative hypothesis is this: it would seem, intuitively, that a change of parametric value in the upward (increasing) direction might produce a greater change in subjective intensity than would a corresponding decrease in parametric value. Thus, such a simple alternation between equal increasing and decreasing parametric intervals as that shown in Figure 8 might really be responded to as though it were something like the plot of Figure 9, with a greater separation associated with the ascending interval. In this case, the factor of similarity would play a decisive role in the perceptual organization of the series into three sets of two elements—whereas, in the first plot, no influence of the similarity-factor in this particular grouping could have been apparent.

A comparison of the three hypotheses suggested above will reveal the fact that each of them represents the intensity-factor as a special case of either proximity or similarity. This can be taken to mean either that the latter factors are really the more basic—the intensity-factor being reducible to one of these—or alternatively, that the analysis (and thus the analyst) is so biased in favor of the factors of proximity and similarity that a more fundamental aspect of the intensity-factor remains in obscurity. Doubts about this may perhaps be removed in the later course of this book, during which proximity and similarity (and especially the latter) will be found to be of unique significance in the unification of musical forms on all perceptual levels. The grouping force of the accent is limited in its effectiveness to relatively short time-spans—serving primarily to articulate successive clangs or shorter elements of clangs—whereas the factor of similarity produces grouping tendencies throughout much longer periods of time, affecting not only the formation of clangs, but also of sequence, longer sections and even entire pieces. It is for this reason that it has seemed appropriate to distinguish between "primary" and "secondary" factors of cohesion and segregation, as defined earlier.

What has already been said about the uneven distribution of attention in the vertical dimension, produced by differences in intensity among concurrent elements, brings up another point which should be mentioned here, although it is not directly related to the question of unit-formation *per se*. When the attention is focussed upon one element or group of elements more directly than it is upon others in a clang, the relative musical importance of the various elements must obviously be different, with the less intense elements taking a subordinate role in the total configuration. This will still be the case when the intensity-differences are great enough to produce subdivision into two or more concurrent clangs (as long as we are considering only one parameter at a time)—the result being typified (in conventional musical terms) by the distinctions between principal and secondary voices, main melodic part versus accompaniment-figures, etc. It should be evident that such distinctions are generally produced by differentiations in parametric intensity, either by the composer or the performer, or both.

The situation here is analogous in many respects to the distinctions between *figure* and *ground* in visual perception—the figure generally being distinguished by what Koffka calls a greater "energy density," and by a higher degree of "internal articulation" than the ground. [11] The analogy between these characteristics and what I have called parametric intensity is obvious, particularly in view of the generality of the definition of the vertical ordinate of the graphs given earlier ("any distinctive attribute of sound, in terms of which an ordinal scale might be constructed"). Vertical and temporal density have already been mentioned as two such

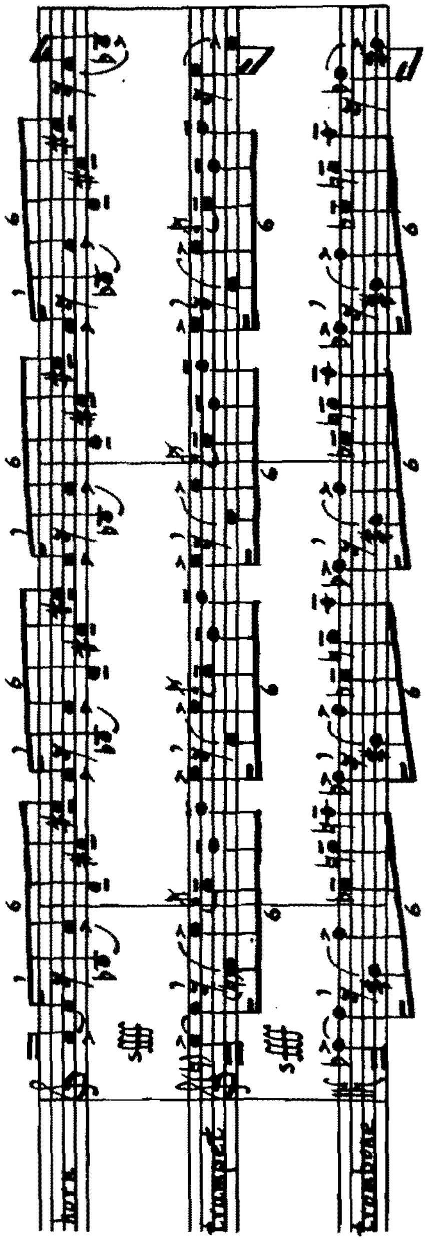
attributes, and the more general notion of *degree of articulation*—the rate of change in parametric values discussed in Section I—can also be considered a parameter to be ordered in a scale of intensity-values like the others.

At this point I want to summarize what has been said so far about the factor of intensity, with respect to both the vertical and the horizontal dimensions of the perceptual model. (1) In a collection of sound-elements, the vertical distribution of attention *at any moment* will be such that, if the differences in the intensity of the various elements are not too great, the more intense elements will tend to be in sharper focus than those of less intensity. On the other hand, if the differences in parametric intensity are considerable, subdivisions (into separate clangs) are likely to arise, as a result of the cohesive and segregative effects of the similarity-factor. (2) In a collection of sound-elements, the temporal distribution of perceptual attention—*from moment to moment*—will be such that, if the differences in the parametric intensity of the elements are considerable, successive clangs will tend to be formed which are *initiated* by the more intense, accented elements.

These two statements might be combined in a general formulation of the *factor of intensity* as follows: *in a collection of sound-elements, among which there are considerable differences in parametric intensity, clangs will tend to be formed in which the more intense elements are (1) the focal points, and (2) the starting-points of these clangs*—other factors being equal.

A fourth factor which can influence clang-formation is the *factor of repetition*. If a repetition of parametric profile is perceived within a series of sound-elements, this alone may produce a subdivision of the whole series into units corresponding to the repeated shape—the perceptual separation between the units occurring at the point just before the first repeated element. That this is a relatively independent factor is indicated by the fact that it can determine perceptual organization even when most of the other factors would tend to produce different groupings, as in example 12.

I am not prepared to offer any explanation of the way in which this factor might function, nor even such hypotheses as were suggested to account for the intensity-factor. It is evident, however, that the factor of repetition involves memory, and more specifically, a process of comparison of what is being heard with what has already been heard. Why this should result in unit-formations in the case of repetition is not so evident. The condition described does suggest,



Example 12. Edgard Varèse, "Octandre," III (mm. 56-58).

however, that there may exist in the listener a positive tendency to group successive sounds into more or less circumscribed units—a tendency, that is, which is independent of, or prior to the objective conditions given in the music. The factors of cohesion and segregation which have been analyzed here would thus turn out to represent not so much active "forces," but rather facilitating conditions—i.e. objective conditions which facilitate the listener's perceptual organization of the sound-elements into clangs. In any case, whether one wishes to consider these factors as causal forces or simply as "facilitating conditions" really makes little difference from a musical point of view—i.e. as long as one's primary interest is in their actual *effects* in musical perception.

We come now to a consideration of those factors of cohesion and segregation which I designated earlier as *objective set* and *subjective set*. The word *set* is used to mean, in general, a prior psychological attitude—involving expectations or anticipations—which may effectively determine or alter the perception of present and future events in the perceptual field. The term *objective set* is borrowed directly from Wertheimer (op. cit.) who used it to describe a factor influencing visual groupings that has an analogous counterpart in musical perception. The term *subjective set* is adopted here as an extension of the implications in the first term, and refers to a whole group of factors such as past experience, learning, habit, association, etc., which Wertheimer mentions, but in a somewhat negative way—because of the overvaluation such subjective factors had received in psychological theories whose basic premises the gestalt psychologists were opposing. The general theoretical situation at the time (1923) involved an active conflict between older, "elementaristic" and "associational" theories of perception and the newer concepts of gestalt psychology, resulting in what may seem to us now to be an undue neglect of such subjective factors in the writings of the gestalt theorists. It is evident now that any really complete evaluation of the various forces involved in musical perception will have to take into account such factors as earlier musical training, cultural orientation, familiarity with the style of the period or of the composer of the work being listened to, etc. And yet one will find a similarly disproportionate treatment of the objective versus the subjective factors in this book, although for somewhat different reasons. Some limitations had to be imposed from the beginning, and I have attempted to restrict my field of inquiry to the more objective side of the musical experience—i.e. to those aspects which may be referred directly to the sounds and sound-configurations which are the materials of the music.

It is quite impossible to make any absolute distinction between the objective and the subjective aspects of the musical experience—and similarly, it is often difficult to decide where to draw the line between the factors of objective set and subjective set, since both of them are "subjective" conditions in some sense, and any distinction we might make would probably seem arbitrary to a psychologist. However, I shall adopt the following heuristic definitions of the two factors, in order to facilitate the analysis, and incidentally to define more explicitly what is to be considered outside of the self-imposed boundaries of the present investigation. *Objective set* will refer to expectations or anticipations arising during a musical experience which are produced by previous events occurring *within the same piece*, while *subjective set* would refer to expectations or anticipations which are the result of experiences *previous* to those that are occasioned by the particular piece of music now being considered. By definition then, objective set should be less variable, from one listener to another, than subjective set, because the former will always have specific analogs or correlates in the musical configurations themselves, while the latter may not.

It will readily be seen that, even after restricting the field to a factor of objective set defined in this way, an enormous number of musical relationships will still be involved. In the most general terms, the factor of objective set will relate to every way in which the perception of an earlier musical event has some effective influence upon the perception of a later event, in a given piece of music. But even within a short composition, such influences are so numerous as to seem virtually infinite to a perceptive listener, and I cannot hope to define or describe completely all of the different forms in which this factor manifests itself. [12] Here I shall mention only three typical ones, with the understanding that there may be others that are just as important to the musical experience.

One of the most common examples of objective set takes a form which might be called *rhythmic inertia*, and is the source of the perception of syncopation, where an accent or metrical impulse is perceived in some way that does not correspond to the actual accentuation in the music at a given point. What seems to be involved here is a psychological or kinesthetic tendency towards rhythmic repetition—the maintenance of a previously established rhythmic structure—which can determine the perceptual organization of a

neutral or ambiguous structure (giving it the form of what has already been heard), or introduce new ambiguities in an otherwise unambiguous structure, thus sometimes causing the rhythmic interpretation of a clang to be very different from what it would be if the clang were heard by itself—out of the particular context.

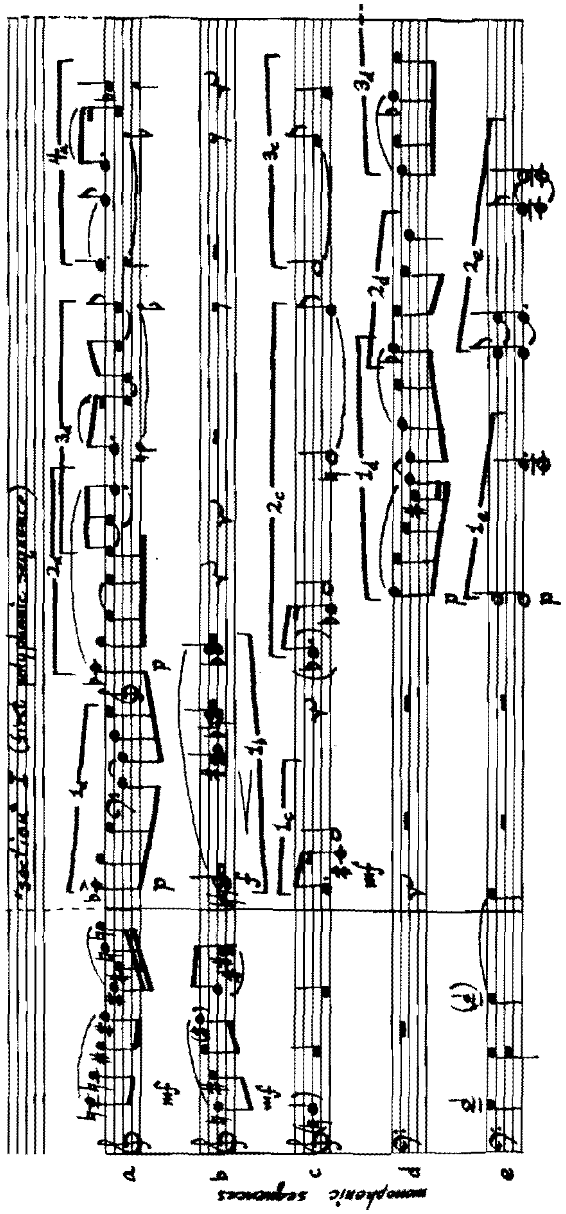
A traditional musical device which takes advantage of this form of objective set is the baroque and classic *hemiola*, in which it may be observed that the subjective rhythmic impulse that is perceived at one moment is a carry-over from the impulse established in preceding measures, and that the new rhythmic structure is often perceived as such a measure or two later than it actually occurs in the music. The strength of such devices depends, as does that of most of the other forms of the factor of objective set, on the establishment of some more or less constant or recurrent condition, and for this reason they are often much less important in 20th-century music than they were in earlier music. But even in 20th-century music, some degree of rhythmic inertia is probably always involved, although its relative effectiveness may be slight by comparison with other factors.

Similar to the above, but not identical to it, is the more general condition whereby the establishment of specific referential norms—whether tonal, metrical, or other—provides a standard of comparison for later events, with more or less specific implications as to the interpretation of these events. Here again, the most obvious examples would come from earlier tonal music—one of the principal characteristics of the traditional tonal system being just this establishment of a referential pitch-level, with respect to which all other pitches receive a specific interpretation. Similarly, when a particular meter is established and maintained throughout a piece of music, or a section of a piece, subsequent events acquire specific rhythmic implications by virtue of their position within that metrical structure (e.g. upbeat vs. downbeat)—the syncopations mentioned above being a special case of such implications. It might be noted here that although it is *objective set* which makes these implications specifiable in the first place, the question as to what particular interpretation will be given to them depends largely on *subjective set*. Thus, for example, the existence of a clear tonal center on C makes the meaning of every other pitch potentially specific, but whether a G is to serve as a "dominant" in that context depends on other factors that include musical conventions that have been learned.

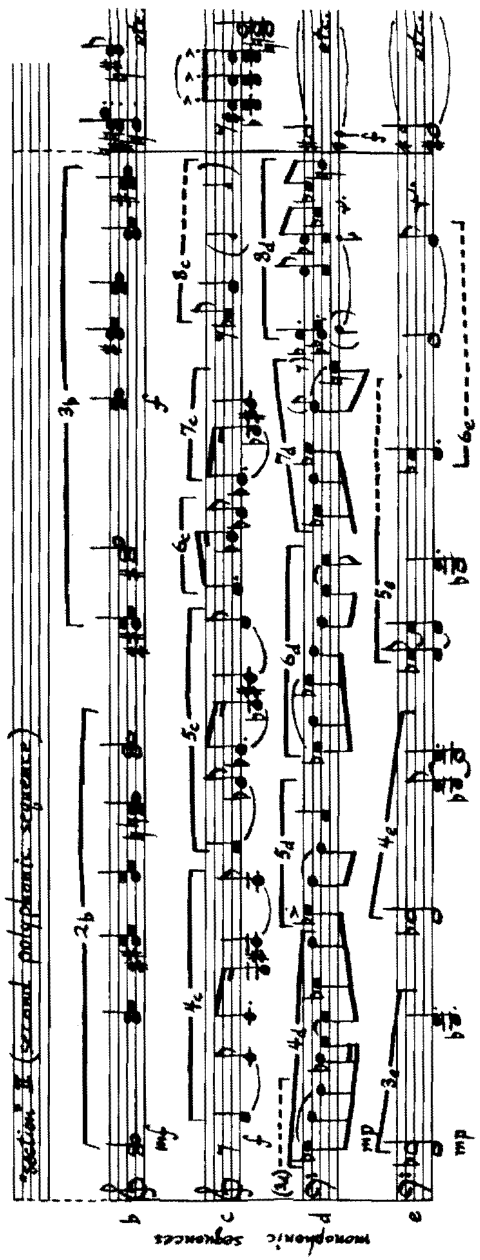
Again it may be said that the importance of objective set has diminished in 20th-century music, but that it must still be present, if only on a smaller scale. That is, the very perception of pitch-intervals represents a sort of primitive form of the same factor. At the lowest level of the perceptual timescale, each sound represents a "referential norm" with respect to the sound that follows it, so that the conditions of objective set can never really be absent from the musical experience.

The third example of this factor—singularly important in most music, though perhaps somewhat less so now than in the past—involves thematic reference, recurrence or recall. This condition depends, more than do the first two, on the longer-range faculty of memory—and is thus less immediate than the others—but it is also capable of altering or determining the perceptual organization of later configurations which are similar, or otherwise related to configurations that have already been heard. That is, a given configuration may have a very different significance when it is perceived as "a variation of" some earlier one, than when it is heard as an entirely new configuration. And, as with the first two examples of objective set, the best sources of examples of this type of relation will be in pre-20th-century music.

It will be apparent from what has already been said that the more radically the new music departs from the conventions of the tonal system and traditional methods of thematic development, the less active do many of these manifestations of objective set become. It might be noted however, that the 12-tone technique and many of the more recent serial procedures seem to be at least partly motivated by a desire to re-institute the cohesive forces of this factor in some new and different way. This is especially clear in the early propositions of the 12-tone method, where the tone-row is treated both as a thematic entity and as an ever-present referential norm of pitch-interval relations—and thus represents an attempt to combine into one form what had previously been two separate sources of cohesive force. Whether the 12-tone technique does this successfully or not is another question; the point here is that the intention behind it can be understood in this way, and it is quite possible that still other means may be found to restore these forces. On the other hand, it may be that the use of these various forms of the factor of objective set corresponds to a more specialized musical attitude, characteristic of one particular historical and cultural *milieu*, and becoming less and less prevalent in our own time. As I have tried to show here, the factor of objective set is by no means the only powerful force active in the perceptual organization or unification of musical configurations.



Example 13 [part one]. Charles Ives, "Concord" Sonata ("Emerson," p. 3).



Example 13 [part two]. Charles Ives, "Concord" Sonata ("Emerson," p. 3).

About the factor of subjective set, very little will be said here, except to note that there is one class of musical phenomena whose effects are closely related to those of thematic reference described above under the category of objective set, but which result from experiences previous to the piece of music in which they exert their effect. I refer to the use of familiar sounds or sound-configurations in a new context—whether these are in the form of more or less exact quotations, or of more general stylistic features. Typical examples of the former may be found in works by Charles Ives, and of the latter in the music of Berg and Bartók, and it is important to note that such devices can have very powerful structural functions in the articulation of the larger form of a piece of music. In any very long work, thematic references between more remotely separated points in time must partake of some of the characteristics of such references to musical ideas already familiar to the listener, and the distinction between objective and subjective set must be understood to include this region of ambiguity in such cases.

In my remarks about the factor of repetition on page 41 I mentioned that a process of comparison was involved—a comparison of what was being heard at a given moment with what had already been heard. To some extent, the factors of proximity, similarity, and intensity would also involve such comparisons, though in none of these instances is the process necessarily conscious. Now the factors of objective set and subjective set may be said to involve a comparison-process also, but in this case it is of a different sort. These factors depend upon the perceptual comparison of what the listener hears at a given moment to what he *expected* to hear at that moment—rather than simply to what he has *already* heard. Again, the most appropriate theoretical definition of these factors would probably involve the concepts of information theory, and more specifically, the theory of "semantic information" based on "inductive probabilities," proposed by Bar-Hillel and Carnap. [13] Unfortunately, it is not within the scope of the present book to elaborate on these relationships to information theory, but I mention them as fruitful possibilities for further investigation.

In order to review some of the principles developed in this section of the book, I have selected for analysis a more extended musical example (ex. 13) in which nearly all of the gestalt-factors of cohesion and segregation may be seen in operation. This passage—taken from the first movement ("Emerson") of Charles Ives' "Concord" sonata for piano—deserves very careful study, because it represents a highly refined application of numerous devices by means of which clangs and sequences may be compositionally organized to achieve a truly polyphonic musical texture. At least two, and more often four separate distinct lines are here developed simultaneously, with a high degree of rhythmic independence (from the standpoint of the phrase-structure—corresponding to the durations of the successsive clangs—delineated within each of the individual sequences). This results in a complex polyrhythm that could never be perceived as such if the several (sequential) lines were not heard as separate strands in the total musical fabric. And this means, of course, that each of these simultaneously-developing sequences must be, in some way, both internally unified, by some cohesive force that connects the successive clangs into one larger configuration, and at the same time, that each sequence must be differentiated from the other sequences by a segregative force that maintains some boundaries between them. It will be instructive to analyze the passage in order to determine specifically how this *polyphonic differentiation* is achieved here—what factors are involved, and in what way they are manifested at any given moment.

In example 13, I have rearranged the notation of the music in such a way that the individual parts can be seen more clearly as separate lines, or what will be called *monophonic sequences*. These will be designated as sequences *a* to *e*, according to their predominant pitch-register—from high to low. The successive clangs in each monophonic sequence are shown bracketed, with Arabic numerals corresponding to their order of occurrence in each sequence. When individual clangs are mentioned in the text, they will be designated by this number, with a subscript to indicate the sequence in which they occur—thus 3a, 5c, etc. The passage constitutes two successive *polyphonic sequences*, which will be referred to as "sections" I and II, respectively—their boundaries being given in the example by the three bar-lines (there are no bar-lines in the original notation). The portions of music that precede the first, and follow the third bar-line are shown to help illustrate certain observations that will be made about aspects of one's perception of the main body of the example that are influenced by conditions outside of it—i.e. in connection with objective set and subjective set.

It should be noted first that the factor of proximity can have very little influence in the polyphonic differentiation of the several monophonic sequences in an example of this kind. Polyphony involves the independent development of *simultaneous* parts,

whereas the effect of the proximity-factor is to neutralize the independence of simultaneous parts—to "fuse" them into a single gestalt. Thus, polyphony is only possible when other factors are made to function in opposition to the factor of proximity. Within each of the individual monophonic sequences, however, the proximity-factor may be involved in the articulation of the boundaries of successive clangs, as it is in this example, between clangs 1c and 2c, or from clangs 2 to 3, 5 to 6, and 6 to 7 in sequence *d*.

The most effective factor in the creation of polyphonic differentiation in a passage like this is of course the factor of similarity. The internal coherence of sequences *a*, *b* and *c*, at the beginning of section I, is the result, in each case, of a characteristic *loudness (piano, forte* and *mezzo-forte*, respectively), *vertical density* (single tone, tone-cluster, single tone), and—to a lesser extent perhaps, *temporal density*. Conversely, the three sequences are "maintained in relative insulation" from one another by their differences with respect to these same parameters. It is noteworthy that the parameter in which the similarity factor manifests itself here is *not* pitch. Indeed, if sequence *a* had been marked *forte*, or had comprised tone-clusters, the pitch-differences between sequences *a* and *b* would not be sufficient to distinguish the two lines—their elements would be perceived as parts of one clang, at any given moment, rather than two distinct clangs. The C# in 1b, for example, would then be heard as a continuation of the melodic movement at the beginning of 1a (i.e., one would hear Ab - G - E - C#. . . instead of Ab - G - E - (low) D. . ., etc.)—rather than as part of a clang beginning with B (the upper tone of the first element in 1b), as it is now perceived.

But after the entrance of sequences *d* and *e*, similarity of pitch-register becomes much more important as a factor of cohesion and segregation in the music. From that point on, each sequence remains within a relatively circumscribed range and register of the pitch-compass, and this is an effective determinant of both their internal coherence and their mutual separation.

But loudness and temporal density still remain important factors. Differentiation in the latter parameter is the primary source of the separation between sequences *d* and *e*, and if the distinction between the *mezzo-forte* of sequence *c* and the *piano* of sequence *a* is not maintained in the performance of the latter half of section I, these two lines will surely fuse into one (as shown by the smaller notes in the notation of *a* at this point). The same general relationships can be seen to apply to the remainder of the example, where parametric similarities always constitute the primary cohesive force within each of the monophonic sequences, parametric dissimilarities being the primary segregative force exerted between them.

The factor of similarity is thus by far the most important factor in the vertical articulation of the passage into separate linear parts, and yet it is of almost no importance at all in the horizontal organization—i.e. the temporal articulation of successive clangs within any one sequence. It has already been mentioned that the proximity-factor plays a part in this temporal articulation, but much more important in this respect are the other factors—intensity, repetition, and objective set.

The factors of intensity and repetition usually function cooperatively in this example. That is, the temporal boundaries defined by these two factors are nearly always congruent or synchronous—as at the beginnings of clangs 2a and 4a, clang 2c (by a repetition of the rhythmic pattern, dotted-8th—16th—halfnote), and 5c, and finally, in clangs 2 and 5 of sequence *d*. In clangs 6d and 7d, on the other hand, the factors of intensity and repetition may be seen to function independently—non-congruently—with the predominant grouping being determined by the repetition-factor (in cooperation with the factor of proximity, already mentioned as influential at these points).

Objective set is involved in the perceptual organization of this passage in two ways. That is, it influences the grouping of both melodic and rhythmic structures. The previous occurrence of the descending melodic pattern, minor second - minor third - major second—as shown in the introductory measure (the part that precedes the first bar-line)—facilitates the perceptual integration of the low D in clang 1a with the three preceding tones in the higher register (and thus, in cooperation with the similarity-factor as it is manifested in the two parameters, loudness and vertical density, but in opposition to the pitch-dissimilarities that would tend to separate these elements).

In the form of rhythmic inertia, the factor of objective set is clearly involved in many of the metrical ambiguities in this passage. A temporal progression in quarter-notes has already been firmly established in earlier passages, and this pulse is maintained consistently only in sequence *b*, so that the groupings of five and seven 8th-notes in duration, which occur frequently in the other sequences, create a complex polyrhythmic relationship between the several lines.

Thus, five of the six gestalt-factors of cohesion and segregation are more or less actively involved in the perceptual organization of this one passage—with each of the factors of similarity, intensity, repetition, and even objective set, being manifested in two or more parameters. The only parameters that are not involved in this example are time-envelope (since a *legato* technique is the only manner of playing that is appropriate here—there are no *staccato*-indications), and—for obvious reasons—timbre. It is likely that some of the differentiations intended here might have been more easily realized in an orchestral or other medium in which a diversification of timbres is possible. And yet Ives has achieved an amazingly high degree of polyphonic differentiation here without this resource—almost in spite of the medium.

The factor of subjective set has not been mentioned in the foregoing analysis, since it does not play any apparent part in the perceptual organization of these sequences. But I have included—at the very end of the example—the beginning portion of the sequence which follows the passage we have been considering, because it shows one of the versions of the opening motive from Beethoven's 5th Symphony that is used in one form or another thoughout the entire "Concord" sonata. And while it cannot be said that subjective set modifies the interpretation of the clangs at this point in the music, there are many other places in the piece where the listener's familiarity with the motive does make his perceptual organization of a clang or sequence somewhat different than it would be otherwise (that is, if the only factors involved were the more objective ones). I mention this only as a reminder that musical configurations may not always be so amenable to an analysis in terms of such objective factors as have been shown to be responsible for the perceptual organization of this particular example.

In answer to the questions put at the end of Section I, six gestalt-factors have been found to be operative in the unification and segregation of clangs, and in the perceptual organization of musical configurations in general. These are the two primary factors of *proximity* and *similarity*, and the four secondary factors of *intensity*, *repetition*, *objective set* and *subjective set*. One or more of these factors will be decisive in the delineation of the boundaries of any clang or sequence, and the composer—whether he does so consciously or not—must inevitably bring these factors into play in the organization of his sound-materials. It can surely be no disadvantage to him to be able to exert that "conscious control over the new means and forms" which Schoenberg held to be the desire of every artist. And I believe that a more explicit awareness of the gestalt-factors of cohesion and segregation outlined in this section of the book might go a long way toward the formulation of a meaningful and realistic technical basis for such compositional controls. An understanding of these cohesive factors is only a beginning, however, and in the next section I shall try to carry the clang-concept a few steps further—into the realm of musical form.

Section III.

"Then—said Stephen—you pass from point to point, led by its formal lines; you apprehend it as balanced part against part within its limits; you feel the rhythm of its structure. In other words, the synthesis of immediate perception is followed by the analysis of apprehension. . . . You apprehend it as complex, multiple, divisible, separable, made up of its parts, the result of its parts and their sum, harmonious."

James Joyce, *A Portrait of the Artist as a Young Man*, p. 212.

"Now the state, including the shape or form, of a portion of matter is the resultant of a number of forces, which represent or symbolize the manifestations of various kinds of energy; and it is obvious, accordingly, that a great part of physical science must be understood or taken for granted as the necessary preliminary to the discussion on which we are engaged. But we may at least try to indicate, very briefly, the nature of the principal forces and the principal properties of matter with which our subject obliges us to deal."

D'Arcy Wentworth Thompson, *On Growth and Form*, pp. 16-17.

"It is certain that this aspect of pure theater, this physics of absolute gesture which is the idea itself and which transforms the mind's conceptions into events perceptible through the labyrinths and fibrous interlacings of matter, gives a new idea of what belongs by nature to the domain of forms and manifested matter."

Antonin Artaud, "On the Balinese Theater," in *The Theater and Its Double*, p. 62.

Section III.
Formal Factors in the Clang and Sequence.

The proposed foundation for a new conceptual framework for musical description and analysis has been based on the premise that musical perception is organized in terms of aural gestalten of great variety and potential complexity, and that the question of musical coherence and formal "continuity" must inevitably revolve around the more basic question as to the essential factors responsible for the perceptual organization of any musical configuration—any clang or sequence. A first step was taken in the preceding section by isolating these factors and defining the specific conditions that lead to unification and relative segregation of musical gestalten in general—but this is only a first step. The description of a piece of music must do more than simply draw the "bounding-lines" around successive clangs and sequences. We will want to be able to describe the characteristic features of the clangs and sequences thus delimited, and—more specifically—those features which are in one way or another essential to the development of the music, and to the musical experience itself. This means that our concern must ultimately be with musical *form*, in all its multifarious aspects, and at all relevant perceptual levels or temporal scales. But in order to describe the form of a given configuration, it will be necessary to take into account certain other attributes of the component materials of the configuration—attributes which are not strictly "formal," but pertain rather to some general condition or state of these component materials. I shall refer to such non-formal aspects of the sounds or sound-configurations as *statistical* features, and to their formal characteristics as *morphological* features—postponing for the moment any more specific definition or justification of these terms.

Consider first what is meant when we speak of the *form* of any sound or sound-configuration. In musical discussions the word is sometimes used to mean something which would more properly be termed "formal unity," or coherence, and is said to depend on such devices as repetition, recapitulation, "return," etc. But this is a highly specialized, and I think misleading use of the word. The devices mentioned above are means toward the unification of a piece of music, or a section or part of it—they do not in themselves give it its form. They are, in fact, large-scale manifestations of the factor of similarity—or a kind of attenuated form of the factor of objective set—both defined in Section II as factors of cohesion and segregation. But although the very existence of a formal unit or gestalt is obviously contingent upon the existence of *unity*—and therefore presupposes the operation of some cohesive factor—this unity is not synonymous with the actual *form* of the gestalt thus produced.

A second use of the word that is, again, often encountered in musical discussions, is illustrated by such terms as "sonata-form," "ABA-form," "rondo-form," etc., which refer to specific *formal types*, generally associated with particular styles or historical periods. And although each of these formal types may be characterized by certain intrinsic formal features, common to all examples of the type, and constituting the original basis for classification, they tend to represent, in each case, not so much a form, but a *formula*, and are not, therefore, relevant to the problems I am concerned with here.

I shall not, then, use the word *form* in this book in either of the above ways. That is, it will be used neither as a substitute for *unity* or *coherence* (which ought to be designated as such in any case), nor in the sense of "a form" or *formal type*, whether classified or not. The word has another, much more general connotation which is consistent with the meaning it has in other (i.e. extra-musical) fields—namely, *shape* or *structure*—and it is in this sense that it will be used in the discussion of musical form which follows—never forgetting, however, that the application of a concept borrowed from other realms of experience may be no more than a useful analogy, with all the dangers which attend any process of extrapolation from one field to another.

I shall follow the analogy one step further, however, and note that, according to the most common definitions of the terms *shape* and *structure*, the former generally implies a more superficial (i.e. pertaining to "surface") or *external* aspect of form (relating to profile or contour), while the latter (structure) usually refers more to an *internal* aspect—"connections" or interrelations among component parts which (interrelations) are not necessarily apparent "on the surface" of the form—i.e. in its *shape*.

I invoke such standard definitions merely to serve as a starting-point in the task of clarification of terms which must precede any adequate analysis of the problem of musical form. But they are, at best, of only limited use to us, because they relate more to the visual and intellectual "fields of perception" than to the aural. What must be done now is to discover what these terms may actually mean in musical perception. That is, how are shape and structure manifested in the clang or sequence, and in our perception of such configurations. And to begin with we must ask what happens when we transpose these concepts from realms whose primary dimensions are *spatial* into a realm which is essentially *temporal*. The following observations on temporal structure will easily be seen to apply as well to temporal shape, and thus to temporal form in general.

I have defined *structure* as involving the "interrelations among component parts," so that the existence of structure in the first place is contingent upon the existence of subordinate parts within a given gestalt. But even at the most immediate perceptual level, a thing can be resolved into parts only when there are *differences* of some kind between one point or region in the perceptual field and another. [14] For a structure which is perceived in *time*, this will mean differences between one *moment* and another—changes in some attribute of sound from one moment to the next in time. It should be evident that, unless such changes occur within a clang, no "subordinate parts" (i.e. successively articulated elements) will be perceived, and that if no parts are perceived, there can be no "interrelation of parts," and thus, no *structure*—in the sense defined above. The very existence of structure in a temporal gestalt would depend, therefore, upon changes that occur within its boundaries, and the perception of differences between one part and another which result from these changes.

But although there can be no perceptible parts where there is no change, there *can* be perceptible change without any resultant subdivision into parts—i.e. when all the changes that do occur are *continuous*. And in such situations, though we may not be able to speak of *structure*, as such, we shall still perceive a *form*, which can only be defined in terms of the parametric changes that occur from one moment to the next in time. What we perceive in this case is that other aspect of form—*shape*—whose temporal manifestation is again based on change, the perception of differences, etc., just as with structure, and which we can (to some extent) represent graphically as an "outline" or "profile" of the variations of some parameter with time.

Thus, it is the differences between the successive elements of a clang, (and between the successive clangs of a sequence), which

determine the form of the clang (or sequence)—not the similarities, although the latter usually constitute the primary factor of *cohesion* in the clang or sequence, as was shown in Section II. In the case of a relatively simple clang, the morphological features may be defined in terms of the *parametric intervals* and/or *gradients* [15] between its successive elements, although with more complex clangs, and with sequences, the measure of "perceptible differences" is not so simple, and may involve both the statistical and the morphological features mentioned at the beginning of this section. But it will be seen that, even here, the same basic principle is still applicable—namely, that *the form of a musical configuration is primarily determined by the effective differences between its successive parts.*

An accounting of the number of distinct ways in which two elements of a clang may be perceived as *different* practically amounts to a listing of the various parameters of sound—by the very definition of the word *parameter*, as given, for example, on page 32 in Section II—"any attribute of sound by which we are able. . .to distinguish one sound or sound-configuration from another." The method of graphic representation of parametric profiles used in the last section should therefore be useful to us in analyzing the form of a clang, and perhaps we can learn something about the musical form in general by applying this method to a specific example. Let us consider a very simple clang—that heard at the beginning of Varèse's piece for solo flute, "Density 21.5," shown in musical notation in example 14.

Conventional methods of analysis would note first of all the melodic-harmonic aspects of such a clang, which are so simple in this case that a plot of the pitch-shape hardly seems necessary. Such a plot is shown in Figure 10, however, in order to illustrate some of the observations that will be made later. As is obvious even without the aid of the graph, there is very little pitch-variation within this clang, the range being only a major second, and the changes that do occur are all clustered near the beginning, the rest of the clang appearing quite static—in terms of this pitch-profile.

A more complete description of the clang might refer to its rhythmic characteristics—two short tones followed by one long tone. Whereas in the pitch-shape there were three different levels (E, F, and F#), here there are only two—the short tones both having a duration of 1/16th of a whole note—but the range of variation between the lowest and highest parametric values here is much greater than in the pitch-profile. Still, the clang would appear to be

Example 14. Varèse, "Density 21.5," first clang.

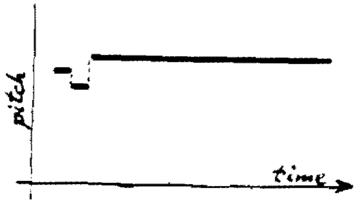


Figure 10.

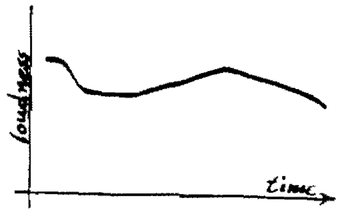


Figure 11.

rather static, the major portion of the clang showing no formal features at all—at least in terms of pitch and duration relations.

But when one listens carefully to a good performance of this piece, the first clang is heard very differently—it has a profile which permeates the whole clang, extending from the beginning to the very end, and giving it a very palpable form which is never static. Obviously, we have still not accounted for the form of this clang as it is actually perceived. And it is probably perfectly evident to the reader that the factor which is responsible for giving shape to the latter portion of the clang—a factor which has been left out of account till now—is the variation in *loudness* that is indicated for the long-held F#. The loudness-profile of this clang might be graphed somewhat as in Figure 11, (where the slight accentuation of the first tone—indicated by the dash under the note in the score—is also represented).

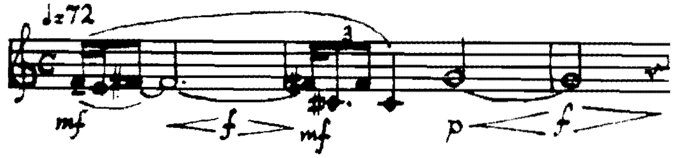
It might be objected here that the fluctuations between *mezzo-forte* and *forte* in this example are only barely perceptible to the ear, or that the extent of dynamic variation is well within the range of "expressive shadings" normally realized by a performer even in the absence of such explicit directions in the score. But this is precisely the important point—that in spite of the small magnitude

of these variations in loudness, the form of the clang as a whole can be profoundly affected by them, acquiring a truly "dynamic" character, a sense of direction, forward impetus, etc., where no other parameter is actively involved. If we are to assume that the perceived form of a clang is a singular, integrated aspect of our apprehension of the clang itself—as I believe we must—we will have to admit that an adequate description of the morphological features of a clang may involve several different parametric profiles—that it will, in fact, involve *every* parameter in which some perceptible change occurs in the course of the clang. And, although it means that our description of a clang's form will not have the singularity—as a description—that is a characteristic of our perception of that form, yet any description will be hopelessly incomplete if it does not at least begin with the simultaneous consideration of all these separate parametric profiles, not just one of them. This does not mean, of course, that all parameters will necessarily be of equal importance in the shaping of a given clang. On the contrary, one of the first things we may discover about the form of a particular clang by such an analysis is which parameter is the most effective in its formation, at any one moment or for the clang as a whole. In the example given previously, the most effective shaping parameter at the beginning of the clang is pitch, but this is clearly not so in the remainder of the clang, where loudness becomes the shaping parameter.

When the formal determinant shifts in this way from one parameter to another within a clang it becomes especially imperative that more than one parametric shape be included in the description of the clang. And this is true not only when we are concerned with the "total form" of the clang as it might be perceived, but also when our interest is centered on one aspect of that form, such as, for example, the *rhythm* of the clang. Here a distinction must be made between what I shall call the *explicit rhythm* of the clang—associated with the relative durations of distinct elements (whose boundaries are delineated by discrete changes in parametric values)—and an *implicit rhythm*, which is determined by the durations from one "peak" to another in the various parametric contours of the clang. When the formative parameter in one part of a clang is not the same as that in another part—as is the case in the Varèse example (where first it is pitch, then loudness)—either the explicit or implicit rhythm of the clang, or both, may become apparent only by means of the simultaneous comparison of the several parametric shapes involved. This is done in Figure 12, where the pitch- and loudness-plots are arranged one above the other with parallel time-axes, for convenient comparison. In addition to the accentuation at the beginning, the implicit rhythm of this clang includes a loudness peak (i.e. a point of highest intensity in that parameter), occurring about half-way through the sustained F#. If properly played, one should hear some degree of rhythmic impulse at that point, even though there is no break in the continuity of this element.

Another example of implicit rhythm, though it involves only one parameter, is the third clang of the same piece (see example 15). Here again, there is an internal impulse to be heard in the clang—a characteristic implicit rhythm—even though the clang consists merely of a single tone, a continuous *crescendo-diminuendo* being its only "articulate" shape and form.

We have so far dealt with an example in which the determination of formal profile shifts from one parameter to another within the same clang. In many clangs this form-determining function is



Example 15. Varèse, "Density 21.5," first sequence.

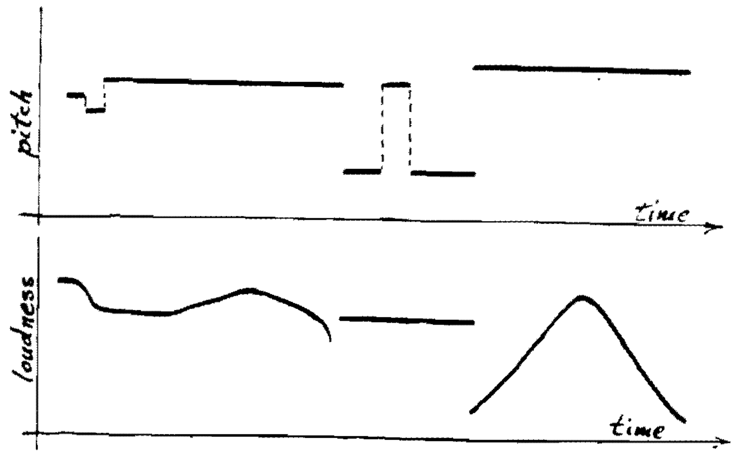


Figure 12.

given to one parameter only, and it is possible to speak then of a primary formal determinant—or *formative parameter*—for the clang as a whole. This will generally be the parameter which shows the greatest amount of variation within the clang—the fastest rate of change—although other, contextual factors may exert an influence which modifies the relative effectiveness of the various parametric shapes from the standpoint of the actual musical impression of clang-form. The thing to be noted here especially however, is that *any* parameter may function as the primary formal determinant in a clang, given certain conditions which may be illustrated by example 15—the whole first *sequence* of the Varèse piece from which the previous example was taken.

Without resorting to the graphic representation used before, it should be evident that these three clangs represent three different situations with respect to the question of parametric determination of formal profile. The formative parameter in the second clang is clearly pitch, since there is no effective change in dynamic level, and very little variation in element-durations (yielding a relatively flat [explicit] rhythmic shape, in addition to the neutral loudness-profile). In clang 3, the determinant of shape is obviously loudness, since there is no variation whatsoever in either of the other parameters, and the objection that might have been raised against my interpretation of the first clang can hardly be maintained in this case. The importance here of the loudness-profile cannot be ignored, not only because the other parameters are constant (or nearly so), but because the variation in dynamic level covers a major portion of the total range of possibilities in that parameter—from *piano* to *forte*—and is no longer commensurate with the ordinary "expressive shadings" of a performer.

The observations that have been made so far in reference to the formal factors at work in the Varèse example relate specifically to shape or profile—and thus to only one of the two aspects of form involved in our initial definition. That is, nothing has been said about *structure*. But it can easily be shown that the same principles apply to structure that have been deduced for shape—i.e. all parameters may be involved in the determination of structure in a musical configuration. Thus, in describing the structure of the Varèse sequence, we would have to note the obvious similarity-relations between the third clang and the second part of the first clang, with respect to dynamic shape (*crescendo-diminuendo*), duration (both being long, sustained), and pitch-region (comprising a half-step relation, which is clearly heard as a melodic movement in itself, bridging the gap created by the C#'s in the second clang).

The conclusions to be drawn from the foregoing are inescapable. Not only is it necessary to include all parameters in any adequate description of clang-form: in addition, we must assume that *any parameter may function as the primary determinant of form in a clang*—if only because it is possible to reduce to zero the degree of articulation of every other parameter within the clang.

In Section II, I tried to show the great functional importance of *similarity* as a factor of cohesion within a clang or sequence. In most musical configurations, one or more parameters change relatively little, within the boundaries of any one configuration, and it is these parameters which do *not* change that give the clang or sequence its unity and singularity—the duration of this relative constancy in these cohesive parameters actually establishing the boundaries of each gestalt. If we compare this with the observations that have been made about the determination of *form* in a clang or sequence, some very interesting relationships become apparent. I have said that the formative parameter in a configuration is usually the parameter which changes the most—exhibits the fastest rate of change—so that it can hardly be, at the same time, the parameter which unifies the configuration because of a relatively constancy of values. That is, *the formative parameter in a given configuration is generally distinct from the cohesive parameter in that same configuration.*

Furthermore, since the morphological outline of a sequence is determined by parametric differences between the successive clangs in that sequence, a rather surprising relationship emerges between parametric functions in a sequence and in its component clangs. That is, *the determinant of morphological outline in the clang will usually be a different parameter from the one which determines the morphological outline of the sequence of which that clang is a constituent part.* This follows from the principle formulated in the previous paragraph, if the latter is combined with certain other principles developed in Section II. There, it was shown that the unity and singularity of a given clang necessarily implied the relative segregation of that clang from others adjacent to it in time, and that these two functions (i.e. unification and segregation) are usually served by one and the same parameter—similarities in that parameter providing the force of internal coherence *within* the clang, and

dissimilarities (in the same parameter) creating the points of division *between* successive clangs. Thus, the differences between clangs, which determine the morphological outline of the sequence, will generally be manifested in the same parameter that serves as the determinant of *cohesion* within each individual clang. And, since the determinant of cohesion—or cohesive parameter—within each clang must be (according to the first of the two principles stated on the previous page) a different parameter from the one which serves as the determinant of form in the same clang, the formal determinant for the sequence as a whole is not likely to be the same parameter that determines the form of each of its component clangs.

Of course, all of the above remarks apply only to clangs and sequences in which the primary factor of cohesion and segregation is the *factor of similarity*. Thus, they would not apply to cases in which the clangs were organized mainly by the factors of proximity, intensity, repetition, objective set or subjective set.

Finally, one obvious exception to these principles must be mentioned. This is the case in which the formal determinant in each clang of a sequence is pitch, but the range of variation within each clang is limited enough to allow for effective changes of register from clang to clang, the shape of the sequence being thereby determined by these changes of pitch-register. But this is only possible because the total potential range of perceptibly different values in this parameter is very great—greater perhaps than in any other parameter—and, in any case, it can only happen when the range of variation within each clang is relatively circumscribed. The more extensive the range covered within each clang in the sequence, the less perceptible will such changes of register from clang to clang become, until pitch is no longer an effective parameter in the process of formal determination at the level of the sequence.

I have repeatedly stressed the fact that the form of a configuration on one perceptual level is the result of changes or differentiations of some kind from one element (or smaller component) to the next within the configuration, because it is of very general significance in the definition of form at any level—not just at the level of the clang—and is manifested in ways that may not be obvious in the more limited discussion of clang-form. For in the first place, only by defining the form of a configuration in terms of parametric intervals and gradients, rather than parametric values themselves, can we account for the phenomenon of *transposability*, which is a unique characteristic of perceptual forms in general and of sound-forms in particular. With respect to the pitch-parameter at least, it is evident that a clang can maintain its morphological identity after transposition—even though the original and the transposed versions have no single element in common. [16] Similarly, within a certain limiting range at least, rhythmic shapes are subject to "transpositions"—i.e. augmentations and diminutions—in which only the relative *proportions* between the parametric values are maintained, not the values themselves (i.e. the element-durations). And I think it possible that such morphological invariance or recognizability after transposition might be found to hold for the other parameters as well, given as great a precision of control over these parameters as we have had in the past over pitch and duration (a precision only recently made possible for these other parameters by developments in the electronic means for generating and recording sounds), and a reasonable amount of time for our perceptive faculties to be conditioned to such relationships.

I do not suggest here that it will ever be possible to perceive *precise* differences or exact "proportions" between loudness- or timbre-levels. These very concepts may be utterly meaningless, from an *aural* standpoint, since the perception of proportional relations in pitch and rhythm is only possible in that they are *periodic* phenomena. But such precision is not necessary to support my assertion here about the transposability of all parametric profiles—if only one is prepared to include less detailed morphological features within the class of transpositional invariants. For example, the *crescendo-diminuendo*, such as occurred twice in the Varèse sequence (example 15), is a recognizable shape, whether it moves from *ppp* to *p* and back to *ppp*, or from *mf* to *ff* to *mf*; (in the example from the flute piece, an interval-expansion is also involved, in addition to transposition, but the conclusions will be the same in either case). This is surely a manifestation of morphological invariance—just as much as is the recognition of a specific melodic gestalt in different registers. The only really essential difference between the two situations is in the relative range of variation in the two parameters involved—the number of different parametric levels that can be perceived, remembered and correlated in a specific way. But this difference in no way contradicts the general principle suggested earlier—a principle which might be abbreviated: *perceived form a function of perceived differences.*

The definition of form in terms of inter-component differences has a second application that was not explicitly apparent

in the earlier considerations of clang-form. The perception of differences involves a higher-order perceptual process than mere sensation—namely, *comparison*—so that the question as to what factors may be involved in musical form can be translated: what are the essential ways in which we are able to compare two sounds or sound-configurations, either on an immediate perceptual level or on a larger temporal scale, where memory, imagination, reflection, etc. may be at work? When an attempt is made to define the essential morphological characteristics of sequences in these terms, two basic factors are encountered, whereas in the problem of clang-form one factor seems to suffice. One of these factors corresponds very closely to that which is involved in clang-formation. That is, one aspect of sequence-form (the *morphological outline*, already referred to) can be defined in terms of the changes of *parametric state* (i.e. mean parametric levels) and other statistical features from clang to clang, in a way that is quite analogous to the definition of clang-form in terms of the changes in parametric values from element to element. But in the sequence another factor emerges—resulting from the fact that we are able to compare clangs with respect to their *morphological* features, not just their statistical features, and the similarities or differences perceived in this way are an essential aspect of our total impression of form at the sequence-level. I shall return to this in a moment, but first some clarification seems desirable regarding my use of the term *statistical.*

When we speak of the *pitch* of a tone in a piece of music—say, for example, the F# in the first clang of the Varèse flute piece—what is it, objectively, that we are referring to? A physicist might answer that this F# is a vibration with a fundamental frequency of 370 cycles per second. The instrumentalist who plays the piece might say that it is the sound produced by a certain fingering on the flute and a certain tension of the lips, diaphragm, etc., in playing the tone. Obviously, the instrumentalist is not describing the sound itself, but the manner of producing the sound. But neither is the physicist's answer any real description of the sound. If we tell him that "370 cycles per second" is an abstraction, and press him further, he might admit that his answer referred to a measurement he might make with a suitable frequency-counting device, which registers the *average* number of vibrations per second in the signal resulting from such a tone. Minor fluctuations in pitch, such as constitute *vibrato*, small variations in pitch that often occur at the beginning and at the end of a tone (*portamento*), and (as may happen in a tone played by an instrumental or vocal *choir*) vibrations whose frequency is very near, but not identical to that of the average mean frequency—none of these "details" is taken into account in the designation "370 cycles per second," nor is it indicated by the musical notation for "F#," in the score itself.

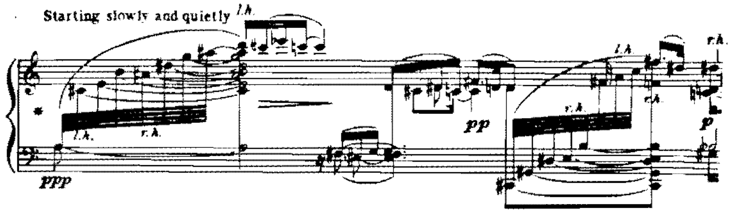
If, now, one looks at the very interesting "performance scores" in Seashore's *Psychology of Music* [17] (pages 35-41, 48-49, 200-203 and 256-272), it becomes clear that the "pitch of a tone" is no simple thing in most music, and can only be defined as some kind of statistical average or mean value of a continuously variable quantity. In these figures it can be seen that the same thing is true of the dynamic level of a tone. And yet we are generally content to represent these variable quantities by a single quantity—a constant—which is nothing but a statistical measure of the sound in some parameter, and we employ this representation both in our notation system and in our verbal descriptions of musical events.

It might be said that we cannot hear these smaller fluctuations in pitch or loudness, but this is manifestly not so. If our listening is such that we do not hear them, it is not because we cannot do so, but rather because *our attention is focussed on a different perceptual level*—a different temporal scale—at which these smaller variations are not relevant in the determination of a parametric profile. Such fluctuations in pitch and loudness influence the timbre or tone-quality of the sound, but they do not affect the pitch- and loudness-contours as such. The latter are determined by the large-scale changes that occur, and are to be defined in terms of the successive values of the averages or means in each parameter. In general then, it may be said that the *morphological* features of a clang will be perceived as a function of the differences between the *statistical* features of its component elements.

I suggest now that this relation between the morphological on one level and the statistical on the next lower level is also applicable to the sequence. That is, the morphological profile of a sequence is primarily determined by certain statistical measures of the clangs in the sequence. These measures would include the changes in *parametric state*, or mean parametric values (pitch-register, mean tempo or temporal density, average dynamic level and vertical density, etc.) from one clang to the next, as well as the total duration of each clang, the extent of the range covered in each parameter, etc. The fact that we have no practical way to measure some of these things precisely is unfortunate, but it in no way

argues against their potential importance for musical analysis, nor their significance in actual musical perception—and this is the most important point, of course. The musical ear *can* "measure" the clangs in this way—and obviously does so—even when the mind of an analyst cannot. [18]

The following example (Example 16) should help to clarify these last remarks. It is the first sequence of the fourth movement ("Thoreau") of Ives' "Concord" sonata, the same work from which example 13 was derived (for the analysis at the end of Section II of this book). The primary determinant of morphological profile in each of these three clangs (indicated again by brackets) is pitch, but how shall we go about describing the profile of the sequence as a whole? Or rather, is there a shape to this sequence that is distinct from the clang-shapes themselves—more than simply the "sum" of these smaller shapes? The changes of pitch-register from the first clang to the second constitute one determining factor that is immediately perceptible when we listen to the sequence—a change



Example 16. Charles Ives, "Concord" Sonata ("Thoreau," beginning).

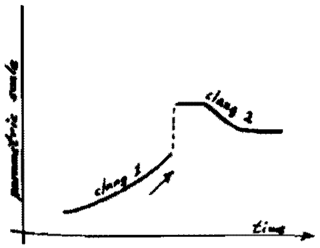


Figure 13.

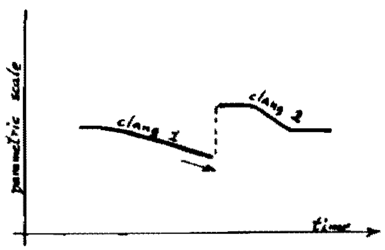


Figure 14.

from a higher register in the first clang to a medium register in the second and third clangs. Another important factor in the shaping of this sequence is the distinction in pitch-*range* or compass between the first and second, and the second and third clangs—first a contraction, then an expansion of range—so that the upper and lower boundaries of pitch in the three clangs describe a movement in the pitch-space even when (as between clangs 2 and 3) an "average" or mean pitch-level might not show any such movement. A secondary determinant of form in this sequence is temporal density, in which parameter the shape of the sequence is represented by the change from faster to slower to faster (i.e. from higher to lower to higher densities) in the three clangs.

Each of these clang-characteristics (namely, pitch-register and range, and temporal density)—in terms of which we are able to compare one clang with another, and thus describe the changes that occur within the sequence, giving it its morphological outline—is clearly a statistical feature of the clangs, and each is a very real aspect of one's immediate and spontaneous perception of the music. Furthermore, it would not be difficult to find examples of sequences in which marked changes in timbre from clang to clang, or in loudness, vertical density, or some other parameter would be the factor responsible for the characteristic profile of the sequence as a whole. Rather than pursue this aspect of the problem any further, however, it should be noted that there is another factor involved in our perception of form in the sequence from the Ives piece—a factor that is quite distinct from, and independent of any of the statistical features of these clangs. Each of the three clangs shows a subdivision into two or three parts, and it can be seen that the second parts of clangs 1 and 2, respectively, are identical in *form*, though they differ considerably in pitch-register. Similarly, the last parts of clangs 2 and 3 are nearly identical in shape, and the first parts of clangs 1 and 3 are quite similar in their general upward motion, if not in the particular interval-relations they involve. These morphological relations (in this case, of identity or similarity) between component clangs (or parts of clangs) in a sequence constitute another important factor in its formal characterization, and must be considered in any satisfactory analysis of sequence-form.

We find, therefore, that the form of a sequence may be conditioned by two distinct and independent factors, which correspond to the two basic ways in which we may perceive differences between clangs—that, is, to the ways in which we can

compare them. Two clangs may be compared with respect to both their statistical and their morphological features, and an adequate description of the form of a sequence may have to include both kinds of differentiation, although one or the other of these might be the more important formal factor in a particular sequence. As for the statistical variations between successive clangs, little more needs to be said, since the same observations that were made about clang-form will also apply to the morphological profile of the sequence. I shall merely repeat here the most basic of the principles established earlier in connection with the clang—that all parameters must be considered, and that any parameter may serve as the primary determinant of form in a musical configuration.

The morphological relations between clangs, mentioned on the previous page, are the source of a kind of formal characterization that is unique to the sequence, since it is not encountered at the level of the clang to any great extent. One can distinguish three basic types of morphological relationship possible between any two clangs: (1) they may be identical, (or nearly, i.e. effectively identical) in form—with respect to one or more parameters; (2) they may be entirely dissimilar and unrelated in form (again—in one or more parameters); and (3) they may be partially similar, or related in form—revealing or implying some kind of morphological *transformation*, by means of which one clang was (or might have been) derived from the other. I shall call the first of these an *isomorphic* relation, the second *heteromorphic*, and the last *metamorphic*—each of these terms being understood to refer to specified parametric shapes, except perhaps in the exceptional cases in which all of the several parametric profiles of the two clangs exhibit the same relation, or in which it is clear that only one parameter is being considered.

These designations can be applied not only to successive clangs, but to any two clangs, regardless of where they happen to occur in a piece of music. In addition to this, they can often be used to characterize a whole sequence, defining what might be called its *morphological type*—whenever the sequence involves internal relationships of one kind consistently. Many sequences, of course, will include more than one type of morphological relation between their component clangs, and these we might call *compound types*—although a meaningful description of this aspect of sequence-form in such cases would still require specification of the particular relations included in that sequence.

In terms of the above definitions, the Baroque sequence would be an *isomorphic* sequence—with respect to pitch, at least. By contrast, most sequences in the early "athematic" music of Schoenberg and Webern are, of course—and by intention—*heteromorphic* in most parameters, though not always. In Schoenberg's piano piece, Opus 11, number 3, for example, the pitch-contours and dynamic shapes are nearly all heteromorphic—throughout the whole piece, not just in one sequence—and yet the rhythmic relations (i.e. the morphological relations between the various profiles of the duration-parameter) are nearly all isomorphic or metamorphic, since they can all be related (by way of various kinds of transformations) to two or at most three "basic shapes" heard in the first few bars of the piece. (See example 17 for the transformations of one of these shapes.)

Finally, it is evident that isomorphic relations with respect to that aspect of the pitch-parameter which is independent of octave-transposition (i.e., *pitch-chroma,* as opposed to the more indefinite *pitch-height*), are bound to occur very often in the systematic 12-tone music of Schoenberg, Webern, Berg and others, although the situation is considerably complicated here by the fact that the actual boundaries of the clang in this music do not necessarily coincide with identical portions (or forms) of the series, (so that it would be quite possible, in 12-tone writing, to avoid isomorphic relations altogether). For the same reason, the *isorhythmic* devices of early Renaissance music *may* result in isomorphic sequences with respect to the duration-parameter, although they need not. Very often they do not do so, and this is simply because the rhythmic patterns do not always coincide with the gestalt groupings (clangs) that are actually perceived, but rather overlap these in various ways.

Isomorphic and heteromorphic relationships represent two extreme poles—two outer limits—of complete similarity and complete dissimilarity between clangs, and it is to be expected that the largest number of actual sequences, and the most commonly occurring morphological relation between clangs, would fall somewhere between these two extremes—within the class of *metamorphic* relations. Different types of metamorphic relation might be defined by reference to the various kinds of morphological transformation which can be applied to a clang, yielding a new and different clang which still bears enough resemblance to the original to be perceived as a variation of the first clang. Such transformations would include, for example, (1) expansions or contractions of

the intervals between the elements of a clang (without altering its essential "topological" features—i.e., the distribution of relative maximum, minimum, and intermediate parametric values in the profile); (2) mirror-forms (inversion, retrogression, and retrograde-inversion) of one or more of the parametric shapes of a clang; (3) clang-extension or compression, by way of (a) the interpolation or elision of elements (i.e. *internal* extension or compression), or (b) the addition or superposition of elements, or the subtraction of elements (i.e. *external* extension or compression); and (4) permuta-

Example 17. Arnold Schoenberg, op.11, #3, transformations of a rhythmic shape.

tions of the vertical order or distribution of concurrent elements, and even perhaps permutations of the temporal order of elements or larger parts within a clang—although this last is not strictly a morphological transformation, unless the parts thus permuted represent substantial and morphologically definitive portions of the original clang, and thus constitute, in themselves, actual clangs.

Examples of such morphological transformations are so numerous in the literature of musical analysis that it should not be necessary to illustrate them here. I have listed them merely to give an indication of the great variety of transformations that may be included in the single category of metamorphic relations—and my list is probably not complete. My primary intention, however, is not to classify, but to *clarify*, and the first step in the direction of clarity is the differentiation of a large field of possibilities into its real and relevant parts—which means here the definition or delineation of all essential and independent factors that may be involved in the larger field of musical form.

There is another side to the relation between the form of a sequence and the forms of its component clangs, which is not yet accounted for by the above definitions of morphological relations and transformations. The perceptual process presupposed there was—as in the case of clang-form—a process of comparison, but it is clearly a rather abstract, intellectual faculty that is involved, one that is dependent upon memory and imagination. For a given clang to be heard as morphologically related to another clang in these terms means that they must both be "present" to the mind in their more-or-less complete forms—i.e., they must already have occurred and passed (become *past*), and must be, at the moment of comparison, stored images, independent of the temporal order in which they originally occurred. This is not, however, the only way in which the form of a clang is perceived, nor is it the only way in which the morphological features of a series of clangs can affect the form of the sequence containing them. This might perhaps be clarified by the following considerations.

The perceived form of a clang must include both a "dynamic" and a "static" aspect, according to whether we view it from the standpoint of the immediate, progressive temporal experience we have of this form, or in terms of the above-mentioned memory-comparison—which is of necessity independent of the original temporal experience. The first is related to one's direct kinesthetic response, always more or less sharply focussed on the immediate

present. Each moment defines only itself, and yet each is continually giving way to the next moment in time. On the other hand, although each momentary event passes away, to be replaced by a new event, those in the past are not thereby lost to us irretrievably. They may be retained and stored in the memory for indefinite periods of time, during which they remain more or less available for comparison with later events—a process which transcends the purely temporal aspects of the original experience. What this amounts to is a kind of "de-temporalization" of the musical images, and—although one should hesitate before calling it therefore a "spatialization" of these images—it has certain features in common with spatial perception. Only in memory can we truly perceive any moderately complex or extended clang "all at once," as a whole—and yet we *are* able to do this in a way that is similar to our perception of visual gestalten. For this reason it does not seem entirely inappropriate to employ such terms as are derived from visual or other realms of experience, such as shape, structure, profile, etc.—so long as we recognize that these represent, at best, merely *one* aspect of our perception of temporal gestalten.

For that other aspect of perceived clang-form which is specifically related to immediate, temporal progression, we need other terms which—although they too may have to be borrowed from extra-musical fields—will at least relate to the dynamic aspect of the musical experience in the same way that *shape* and *structure* relate to the static aspect. For this dynamic characteristic of clang-form, the words *gesture* and *movement* seem appropriate. The concept of clang-form would include, then, both shape and gesture, structure and movement, the static and the dynamic—like positive and negative poles of a descriptive field, neither of which can fully represent the total field, although they are both necessary to any full description.

The relevance of all this to the problem of sequence-form may be illustrated by considering one manifestation of the dynamic aspect of clang-form—namely, the *directionality* implicit in a gesture. A conjunction of two clangs in which their gestural characteristics (symbolized by the arrow under clang 1) are related as in the idealized plot in Figure 13 will have a very different effect on the perceived form of the sequence than would the one shown in Figure 14. In the first case, the direction of movement in clang 1 will considerably mitigate the discontinuity that marks the break between the two clangs, while the effect in the second case will be to emphasize the contrast between the two—even though the differential intervals between the clangs are the same in both instances (as measured from the end of the first clang to the beginning of the second; if mean parametric values are used as a measure, the interval-magnitudes would actually be in an *inverse* relation to the perceived discontinuities). The essential difference between the two situations resides in the relations between the direction of the *gradient* in the first parametric profile (in each example) and the direction of the *interval* between the profiles of clangs 1 and 2. And in general, it can be said that the degree of effective contrast between two clangs (with respect to a given parameter) depends as much upon the direction of the initial gradient as it does upon the magnitude of the interval separating the two clangs. And this "degree of effective contrast" between two successive clangs in a sequence is the proper measure of sequential profile at that point—supplementing or replacing the simpler measure of the change in parametric state.

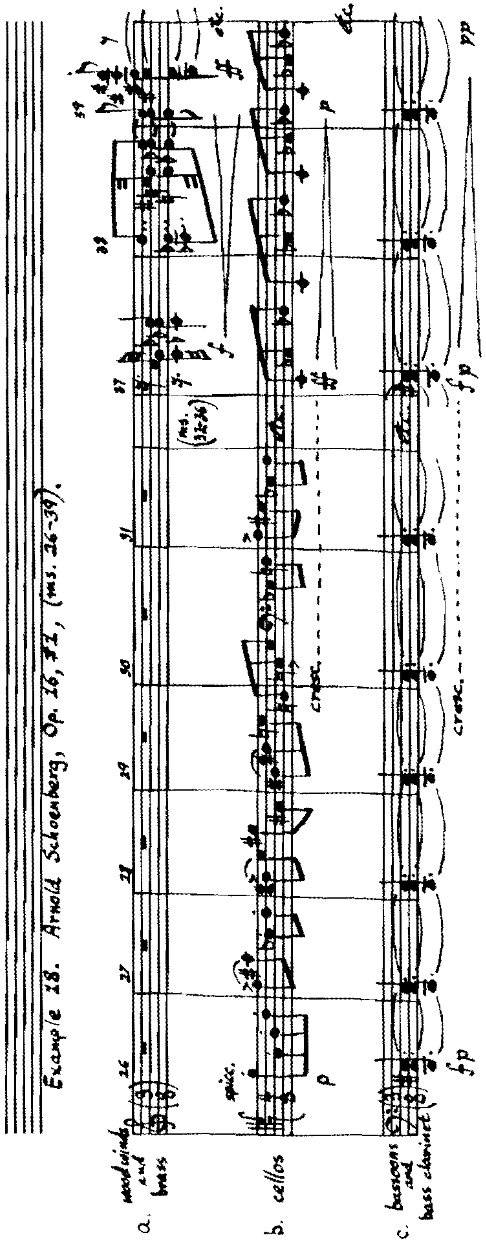
I have related this factor of directionality to the matter of clang-morphology, although it seems also to partake of some of the characteristics of clang-statistics—and here perhaps we have a borderline phenomenon for which my earlier distinctions between the morphological and the statistical begin to break down. However, these distinctions have proved useful up to this point, in helping to uncover several different factors that contribute to the formal characteristics of the sequence, and I see no reason to abandon them because of the appearance of a factor which pertains to both categories. Such dynamic aspects of clang- and sequence-morphology may, in fact, be interpreted as transitional factors, which bridge the gap that would seem to fall between the temporal, more purely sensory aspects of musical perception, and the "de-temporalized," mnemonic, more intellectual aspects that are involved in the musical experience. The hiatus between these two realms, which seems to arise so inevitably in most psychologies and philosophies, is perhaps something that is in the nature of the basic attitude toward experience that is involved in such disciplines, rather than in the nature of experience itself.

Two further distinctions must now be made with regard to the basic types of sequence. The first of these involves the perception of sequences with respect to the time-dimension, the second relating more to the vertical characteristics of sequence-structure. In Section I, the sequence was defined as "a succession of clangs. . .constituting a musical gestalt on a larger perceptual level or temporal scale." Implicit in this definition are (1) some degree of

unity—though the sequence will be less unified than the clang, in perception, and (2) a temporal articulation into distinct parts—the successive clangs—whose own unity and relative segregation within the sequence are determined by the gestalt-factors described in Section II. For the most part, the factors responsible for clang-delineation are "objective," in the sense that they can be referred to perfectly objective characteristics in the music itself. That is, they are not arbitrary, and one could predict with reasonable accuracy just where the boundaries of the clangs will be perceived by most listeners. There are certain significant exceptions to this, however, which I shall call *monomorphic sequences*, and these exceptions constitute a class of musical configurations at this level which must be distinguished from the *polymorphic sequences* we have been dealing with so far.

One of the assumptions which must be made, in any attempt to describe musical organization and perception in terms of the gestalt-concept, is that there are some approximate durational limits beyond which a sound or sound-configuration will no longer be heard as an immediate aural gestalt—that is, it will not be perceived as a *clang*. If the duration of a sound is too short—say, less than one-half second—the sound is not likely to be heard as an individual clang, but will become simply an *element* within a larger clang. [19] Similarly, a sound-configuration lasting longer than a few seconds is likely to be resolved into several shorter clangs by the listener, and so be heard as a *sequence*. These durational limits obviously vary, depending upon such factors as the relative simplicity or complexity of the configurations themselves, and upon all the gestalt-factors discussed in Section II, so that there would be no point in trying to attach any absolute values to the upper and lower boundaries of this range. But it is evident that, variable as they may be, there *are* limiting regions to the range, and these must be recognized in our definitions.

Consider then the following examples, which represent two kinds of *monomorphic sequence*. In the first, example 18, the sound designated as *c* (on the third staff) is maintained so long that it cannot be called simply a clang—though the term *resonant clang* would seem to be an appropriate description of its musical character. Its function, as well as its duration, is commensurate with that of a sequence, shaped only by changes in timbre and loudness (changes in the former parameter only occurring several pages later in the score). It is, of course, a subordinate part of the total musical fabric, but



Example 18. Arnold Schoenberg, op. 16, #1 (mm. 26-39).

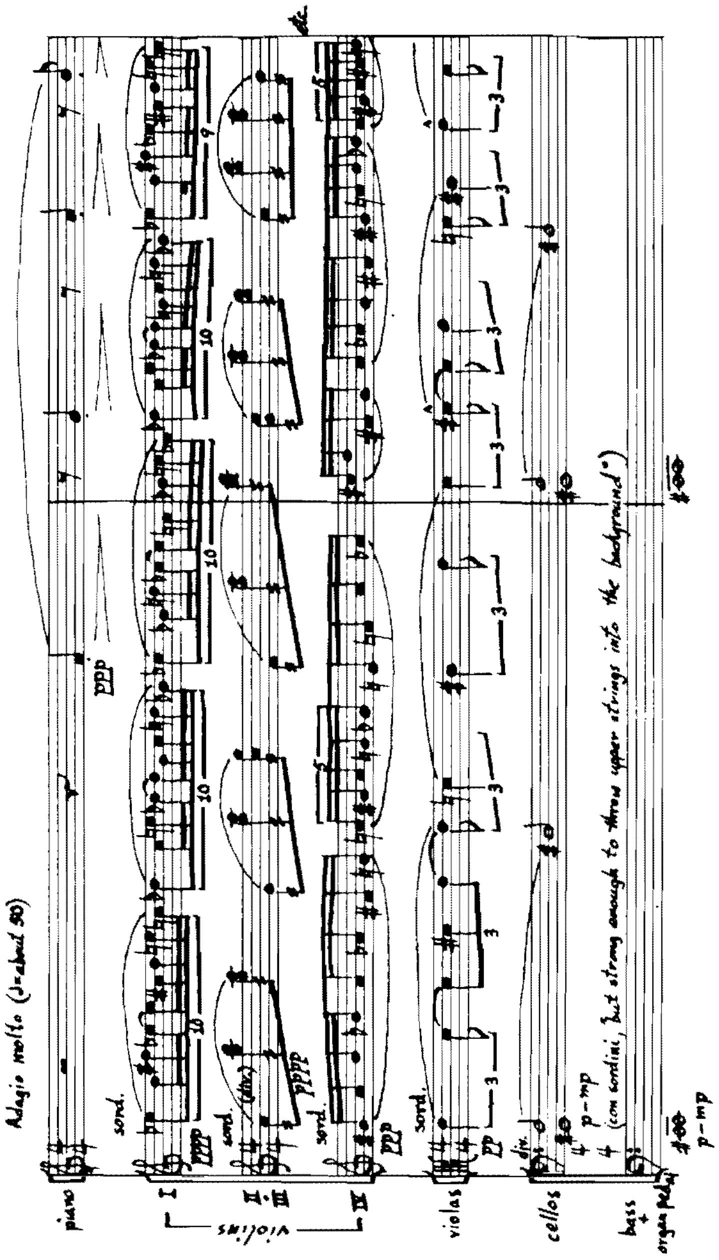
this does not concern us here, since the original definitions of clang and sequence did not involve the question of the relative importance of parts, but simply the delineation of such parts within the texture of a piece of music.

Example 19 shows another kind of monomorphic sequence, in which the changes in sonority are so continuous that the "boundaries" of unit-formations on the order of the clang may occur almost anywhere—i.e., perceptual organization does not seem to be determined by any objective characteristic of the music itself. Yet the configuration is so long that subdivision must occur somewhere, and the groupings that do result will probably be coincident with the rise and fall of each listener's acuity of attention. The musical structure of such sequences is as though composed of an extended succession of elements, rather than a succession of clangs, though this is no more than a very imprecise way of describing the process, and does not apply to the type of monomorphic sequence which results from clang-resonance, as in example 18.

In any case, both the Ives and the Schoenberg examples have this much in common at least: they are extended sound-configurations, of the durational order of the sequence, in which any perceptual grouping or subdivision into clang-like units is almost entirely arbitrary or subjective, not depending upon any clear-cut objective characteristics of the configurations themselves. This last statement may be taken as the definition of *monomorphic sequence*—a type of configuration to be considered as an exceptional or special case of the more general class of sequences. The typical case, on the other hand, would be the *polymorphic sequence*, and the definition of *sequence* given in Section I should be understood to apply only to the latter type.

Obviously, the form of a monomorphic sequence will not involve the morphological relations between component clangs described earlier—but such a sequence will still have an overall *morphological outline* or *profile*, determined by the changes in parametric values from one moment (or element) to the next in the sequence.

The second distinction with respect to type and function at the sequence-level has already been made or implied in an earlier part of this book—during the analysis of the Ives passage (example 13) at the end of Section II. There a distinction between *monophonic* and *polyphonic sequences* was employed in the discussion, though I did not give any explicit definitions of the terms, assuming that the



Example 19. Charles Ives, "Three Places in New England" (III; "The Housatonic at Stockbridge").

intended meanings could easily be deduced from the musical example itself. Here I shall try to define these two terms in a way that is consistent with my earlier usage of them, and it will be seen that I interpret them somewhat more broadly than is common in traditional music theory.

By *monophonic sequence* I mean one in which the clangs are perceived one at a time—even when successive clangs are not simply connected end-to-end, but are dovetailed or overlapped to some extent. In a monophonic sequence, such overlapping connections between clangs serve primarily to provide greater continuity to the configuration—to mitigate the otherwise mechanical effect of simple juxtaposition. The sequence is still monophonic, however, so long as the attention is directed essentially to one clang at any given moment.

But if the degree of overlapping of the component clangs is increased to the point where the sequence is no longer heard in this singular way—the attention now being divided or distributed among two or more clangs simultaneously at certain moments—then the sequence becomes *polyphonic*, as in the Ives example studied at the end of Section II, or the last Schoenberg passage shown (in example 18), where three distinct strata are sometimes sounding simultaneously. It is not simply a question of increased complexity of the sound-materials that is involved here, but rather the use of certain techniques of polyphonic differentiation of these materials—by way of the same gestalt-factors of cohesion and segregation described in Section II.

A truly polyphonic situation is not necessarily created by the addition of new parts to a texture, because these may simply be absorbed by the others in a succession of clangs that become more and more complex, but no less singular. There must be strong differentiations among the various parts for a polyphonic texture to be perceived as such, and since the factor of proximity can play no role here (polyphony implies an independence of parts sounding simultaneously, as was noted earlier)—the *factor of similarity* is virtually the only one that can effect such polyphonic differentiations. That is, there must be clearly perceptible parametric differences between the individual monophonic sequences, and a relatively high degree of parametric similarity *within* each one, before the sequence as a whole can be heard polyphonically.

Twentieth-century music furnishes many examples of this kind of complex polyphony—a polyphony in which each of the individual lines (i.e., monophonic sequences) is itself complex, by comparison with earlier music. And yet polyphonic sequences are not to be found quite as easily as one might imagine, considering the prevalence of more complex textures in the music of our time. Ostensibly polyphonic music is often quite monophonic in effect, in spite of its complexity—or, as it sometimes appears, because of it—since what one actually perceives in listening to the music is

Example 20. Anton Webern, op. 5, #1 (beginning).

essentially a succession of single clangs, some more complex (in their vertical structure) than others, but one at a time, nevertheless—as in example 20. Here, the new parts introduced in contrapuntal imitation (in measures 3 and 4) are not likely to be apprehended as distinct clangs. Rather, what will be perceived, at each of these entrances, is simply an intensification—by means of an increase in vertical density—of the sonority of a single clang.

I do not mean to imply here that such monophony is undesirable, nor even that polyphony as I have defined it is desirable or necessary in music, but simply that one should be prepared to distinguish the one from the other in a way that is more consistent with actual musical experience. I do believe, however, that the developments of a higher-order polyphony of the kind I have been describing constitute one of the most significant characteristics of early 20th-century music, and that the almost limitless possibilites for further development in this direction represent one of the most exciting aspects of music in our own time—mid-20th-century.

Unfortunately, a thorough examination of these possibilities would carry me far beyond the limits of this book—as would a more detailed study of many other problems of musical form. A beginning is all that has been attempted here, and a provisional outline of possible solutions to the most immediate problems that arise in the study of form in music. It is probable that many of the most important questions have not even been asked yet, much less answered. And there is no doubt in my mind that some of the ideas presented here will not stand the more severe tests of practical application without at least some modification or revision. It seems to be in the very nature of musical experience to resist our attempts at rationalization, and to contradict our theories.

But the final test of any concept—and the only valid source of any rationale—must be experience itself, and a musical theory that does not maintain a direct and vital connection with musical experience cannot be expected to survive for very long. I only hope that the observations made in this book may prove helpful in clarifying some of the problems which concern the musician of today, and that they will provide a conceptual framework that is sufficient, in breadth and depth, to form the basis for more refined techniques of musical description and analysis—and eventually perhaps, of musical composition itself.

Endnotes

1. Schoenberg, *Style and Idea*, Faber & Faber, London, 1975, pp. 216-17.

2. Note that the parameters listed here are specifically *musical parameters*—attributes of perceived sound which are the "subjective" counterparts of the physical or *acoustic parameters* (frequency, amplitude, wave-form, etc.). The word *parameter*, when used by itself in this way, will always refer to the musical parameter, rather than to the corresponding acoustic parameter.

3. This is not intended to mean that there is *always* a faster rate of change in the music, but rather simply that faster changes can and do often occur.

4. Op. cit., p. 240.

5. Routledge & Kegan Paul Ltd., London, 1962.

6. In "Physical Gestalten," page 17, as now printed in: *A Source Book of Gestalt Psychology*, edited by Willis Ellis, Humanities Press, New York, 1967.

7. Especially his book, *A la Recherche d'une Musique Concrète*, Editions du Seuil, Paris, 1952.

8. Available in English translation in *A Source Book of Gestalt Psychology*, edited by Willis Ellis, Humanities Press, New York, 1967, which also includes some early papers by Köhler that are of interest from a theoretical standpoint.

9. See the listings for these authors in the bibliography on page 96.

10. An ordinal scale represents a "rank ordering" of relative magnitudes of some attribute, an ordering which involves the distinctions "greater than" and "less than" (indicated on the scale by displacements up or down, respectively), but does not purport to show how much greater or how much less one point on the scale may be than another point.

11. Koffka, op. cit., pp. 186-209.

12. The relationships which can be described show characteristics which indicate that some kind of "field-theory" might provide a basis for the definition of the essential features of this factor—more specifically, some of the concepts of the "topological field" introduced into psychology by Köhler, Koffka and Lewin. The concepts of information theory might also provide such a basis, perhaps even in combination with the field-concept, and this could be correlated with the other cohesive factors in ways suggested on page 36, Section II. All this is pure speculation on my part, of course, but it is sometimes meaningful to point out possibilities in the way of larger relationships, even though these have not yet been clearly formulated.

13. For a review of this theory, and of the concepts of information theory in general, see Colin Cherry's *On Human Communication*, The M.I.T. Press, Cambridge, 1978.

14. Cf. the implications of "segregation" in Section II, and the following remarks by Wertheimer (op. cit., p. 88): "When an object appears upon a homogeneous field there must be stimulus differentiation (inhomogeneity) in order that the object may be perceived. A perfectly homogeneous field appears as a total field [*Ganzfeld*], opposing subdivision, disintegration, etc."

15. The term *parametric interval* will be used here to refer to an approximate measure of the difference between two values (in any parameter, not just pitch)—especially when the change from one value to the other is *discontinuous*. A parametric interval would thus be defined by both a relative magnitude and a "sense" or direction—i.e. up or down on that parametric scale. The word *gradient* will refer to *continuous* changes, also specified by both a magnitude—the rate of change or "slope"—and a direction (positive or negative) exhibited by a given segment of a parametric profile.

16. This transposability of a melodic figure was in fact one of the principal attributes of this particular "*Gestaltqualität*" (viz., shape or form) noted in the 1890's by von Ehrenfels, a precursor of Wertheimer and Köhler in the early development of gestalt psychology. For a description of von Ehrenfels' contribution to

gestalt theory, see Köhler's *Introduction to Gestalt Psychology*, pages 102 to 104.

17. Dover Publications, Inc., New York, 1967.

18. Heinrich Schenker's concept of "middleground" (and perhaps also "background") could be considered a special type of morphological outline "at the sequence-level"—involving the pitch-parameter, and representing one of the many possible measures of "statistical differences" between successive musical configurations, which determine the shape of the next larger configuration.

19. If such a sound were separated—by silences—from the sounds that immediately precede and follow it, it might very well be perceived as a complete clang, but in this case the silences must be interpreted as real elements of that clang, so that its actual duration will no longer be outside of the normal range of durations within which aural gestalten can be perceived as such.

Glossary

A review of some of the more important terms and definitions.

CLANG. A sound or sound-configuration which is perceived as a primary musical unit or aural gestalt. The clang-concept constitutes the nucleus and core—in fact, the essential "heart and soul" of the entire "conceptual framework" proposed in this book.

CLANG-RESONANCE. The sustention or repetition of a clang beyond the normal limits of clang-duration (lasting, that is, longer than a few seconds), resulting in one type of monomorphic sequence.

COHESION AND SEGREGATION, GESTALT-FACTORS OF. Forces (or "facilitating conditions") which determine the perceptual organization—i.e., the internal unification and mutual separation—of clangs and sequences. The primary factors are *proximity* and *similarity*; the secondary factors are *intensity*, *repetition*, *objective set* and *subjective set*.

COHESIVE PARAMETER. See DETERMINANT OF COHESION.

DENSITY, TEMPORAL. One of the seven musical parameters most frequently referred to in this book; a measure of the relative speed of parametric alteration in a clang (or sequence), or the number of successive elements distinguishable per unit time.

DENSITY, VERTICAL. The number of simultaneous elements perceptible at a given moment in a clang.

DETERMINANT OF COHESION. The parameter (or parameters) in which the factor of similarity is manifested, in a given clang or sequence; usually the parameter that varies least—maintaining relatively constant parametric values—within the boundaries of the configuration.

DETERMINANT OF FORM. Generally, the parameter (or parameters) undergoing the fastest rate of change—the highest degree of articulation—in a given clang or sequence, being thus the subject of the listener's most direct and acute parametric focus. This form-determining parameter is usually distinct from the determinant of cohesion in the same clang or sequence, since the latter is necessarily constant or nearly so.

Directionality. That aspect of clang- and sequence-morphology relating to a continuous increase or decrease in values in some parameter, yielding an impression of movement up or down in pitch, loudness, tempo, etc.—i.e., on some parametric scale. The term "singular directionality" was also used in Section I to refer to the fact that each parametric scale is assumed to have an implicit and "absolute" upward and downward direction associated with it, corresponding to an increase or a decrease in parametric intensity.

Dynamic. This word has been used here in two different ways. I have sometimes used the term "dynamic level"—instead of "loudness level"—to refer to some value in that parameter, in accord with conventional musical usage. In Section III, however, it is also used in the more general sense, *vis-à-vis* "static," to describe that aspect of musical perception that is immediately bound to the temporal order of the musical experience, thus involving *gesture* and *movement* (as opposed to *shape* and *structure*).

Element. A component part of a clang, which may either be one of several *successive* parts—corresponding to the internal articulation of the clang in time—or one of a number of linear, *concurrent* parts—coextensive with the clang as a whole. Thus, an element might contain smaller elements. In addition, an element is assumed to be an aural unit—as is the clang—the only basic difference between the two being the degree to which an element is"absorbed" into the larger configuration of which it is a part.

Envelope, or **time-envelope.** The shape of the attack and decay forms of a sound, with respect to changes in amplitude. As a musical parameter, however, the perception of the time-envelope of an elementary sound relates to the impression of tone-quality or timbre, more than it does to the loudness-parameter.

Equivalence, and the "**Principle of equivalence.**" These terms are used in Section I, in reference to the "equal potentiality of any sound being used as a basic [or irreducible] element of a musical

idea" (i.e., of a *clang*). It does not mean an equivalence of musical effect or character, but a "material equivalence," in the sense that any sound might occur within a clang, as an element.

Explicit rhythm. The duration-relations within a clang that derive from discrete changes in parametric values from element to element—being measured, therefore, from one *attack* to the next.

Focus, parametric. The directing of the attention toward a particular parameter—generally the parameter with the highest rate of change or degree of articulation within a given clang or sequence.

Focus, textural. The directing of the attention toward a particular (linear) part or element within a clang, (or a particular monophonic sequence within a polyphonic sequence)—usually that element which is the most intense, in one or more parameters.

Form. That aspect of our perception of musical gestalten (whether these be clangs, sequences, or larger configurations) that involves *shape* and *structure*, and *gesture* and *movement*—as its "static" and "dynamic" attributes, respectively. In Section II, the statement is made that "the form of a musical configuration is primarily determined by the effective differences between its successive parts." At the perceptual level of the clang, this means the changes in parametric values from one element to the next. For the sequence, two factors are involved, because "effective differences" between successive clangs may be perceived in two different ways. These are (1) as changes in the statistical features of the clangs, from one to the next, and (2) morphological relations (similarity, partial similarity, and dissimilarity of form) between clangs, yielding in some cases distinct sequence-types.

Formative parameter. See **Determinant of form.**

Gestalt-factors. See **Cohesion and segregation.**

Gradient. An approximate measure of the rate of change of values in some parameter, when the changes are continuous, rather than discrete. A parametric gradient would be specified by a magnitude ("high" or "low") and a direction (positive or negative on the parametric scale).

HETEROMORPHIC RELATION (& SEQUENCE). The morphological relation of complete dissimilarity of form between two clangs. A sequence in which all the clangs were different in form would thus be a HETEROMORPHIC SEQUENCE.

IMPLICIT RHYTHM. The duration-relations within a clang that derive from the impulses created by "peaks" of intensity in the various parametric profiles of that clang. Since these peaks may occur during continuous changes of parametric values—and thus in the "internal" portions of an element, as well as at its beginning (in the attack)—the implicit rhythm of a clang will be a more inclusive attribute than the explicit rhythm, which is measured from one *attack* to the next.

INTENSITY, PARAMETRIC. In each parametric scale (as described and employed in Section II), the higher of two values is assumed to be the one which produces or corresponds to a greater musical or subjective intensity. The measure of relative height on such a scale is then an indication of parametric intensity.

INTENSITY-FACTOR. One of the secondary gestalt-factors of cohesion and segregation described in Section II, referring to the tendency of an accented sound to be heard as the beginning of a grouping. The relative intensities of several concurrent elements in a clang (or of several monophonic sequences in a polyphonic sequence) is also a determinant of textural focus (see page 41, Section II for a more complete statement of the effects of this factor).

INTERVAL. A measure of the difference between two (discrete) values in some parameter—a meaningful concept even when this difference cannot be specified in any precise, quantitative way, but merely in such approximate terms as "large" or "small," "wide" or "narrow," etc. In addition to a magnitude, an interval will also (like the gradient) have a direction (up or down) on the parametric scale.

ISOMORPHIC RELATION (& SEQUENCE). The relation of complete similarity or identity of form between two clangs (with

respect to a given parameter). A sequence in which all the clangs were identical in form would be termed an ISOMORPHIC SEQUENCE.

METAMORPHIC RELATION (& SEQUENCE). The relation of partial similarity of form between two clangs, "revealing or implying some kind of morphological *transformation*, by means of which one clang was (or might have been) derived from the other." A sequence in which all the clangs were interrelated in this way would be a METAMORPHIC SEQUENCE—probably the most frequently occurring sequence-type to be found in music.

MONOMORPHIC SEQUENCE. A special case of sequence-structure which is not perceived as a "succession of clangs" because "any perceptual grouping or subdivision into clang-like units is almost entirely arbitrary and subjective." This type of configuration is often produced by clang resonance, though not always, and it usually plays a secondary role in the musical texture, as an accompaniment or "background."

MONOPHONIC SEQUENCE. A sequence in which the clangs are perceived one at a time.

MORPHOLOGICAL FEATURES. Those aspects of a clang (or sequence)which relate specifically to its form, as distinct from its parametric state, or other statistical features.

MORPHOLOGICAL OUTLINE OR PROFILE. These terms have been used here to refer to that aspect of form which derives from the changes in parametric values from element to element in a clang, or the changes in parametric state from clang to clang in a sequence. It is assumed to be a kind of synthesis of all the various (single) parametric profiles of a clang or sequence, and—for the sequence—is meant to be distinguished from the *morphological type*, which refers to the specifically formal relations between the component clangs.

MORPHOLOGICAL RELATIONS (between clangs), and SEQUENCE-TYPES. General terms that involve the isomorphic, heteromorphic and metamorphic relations between clangs, and the types of sequence-structure that derive from the consistent use of one or another of these relations in a given sequence.

OBJECTIVE SET. One of the secondary gestalt-factors of cohesion and segregation, defined in Section II as "expectations or anticipations arising during a musical experience which are produced by previous events occurring *within the same piece*." One of the most effective manifestations of this factor is in the form of *rhythmic inertia.*

PARAMETER. Any distinctive attribute of sound, in terms of which one (elementary) sound or sound-configuration may be distinguished from another. Seven parameters have been referred to more or less frequently: namely, *pitch, loudness, timbre, duration, temporal density, vertical density,* and *time-envelope.* Although these are the parameters most often involved in musical analysis (as in musical composition), the more generalized definition given above leaves room for others that may be relevant in certain cases, such as pitch-range, degree of parametric articulation, etc. These are all what I have called "musical parameters," to distinguish them from the "acoustic parameters"—frequency, amplitude, wave-form, etc.—which are their physical counterparts and source. When the terms themselves do not imply any distinction between the"objective" and the "subjective" correlates of a parameter (as is the case with "duration," "density," and "time-envelope") it is still the specifically *musical* parameter that is intended—i.e., an attribute that is actually perceived as a part of the musical experience, not simply subject to measurement or abstract determination of some kind.

PARAMETRIC FOCUS. See FOCUS.

PARAMETRIC PROFILE or **SHAPE.** That aspect of the perceived form of a clang or sequence which is the result of the changes in a particular parameter from one moment to the next in time. Also, the graphic representation of these changes, as employed in Section II and Section III.

PARAMETRIC SCALE. An ordinal scale—i.e., one which gives a "rank ordering of relative magnitudes of some attribute [involving] the distinctions 'greater than' and 'less than' (indicated on the scale by displacements up or down, respectively), but does not show how much greater or how much less one point on the scale may be, relative to another point."

PARAMETRIC STATE. An approximate measure of the average or mean value of all those in a parametric profile of a clang. It is thus one of the main statistical features of a clang—changes in parametric state from one clang to the next constituting the basis of the morphological outline of the sequence.

PERCEPTUAL LEVEL. This term has been used synonymously with TEMPORAL SCALE, to refer to distinctions between the gestalt-organization and perception of configurations of the order of a few seconds or less in duration (for the clang), and those that span longer periods of time and must be much less immediately apprehended as gestalten (viz., the sequence, as well as longer sections and even entire pieces)—though they may be apprehended thus nevertheless, if only by way of higher-order intellectual faculties such as memory.

POLYMORPHIC SEQUENCE. The kind of sequence-structure assumed to be "typical," by comparison with the monomorphic sequence (see the definition of SEQUENCE).

POLYPHONIC SEQUENCE. A sequence composed of two or more monophonic sequences. More precisely, a sequence is called polyphonic when "the attention is divided or distributed among two or more clangs simultaneously at certain moments." Thus, the mere existence of two or more instrumental parts in a contrapuntal passage, for example, does not necessarily mean that the passage is polyphonic—by this definition ". . .There must be clearly perceptible parametric differences between the individual monophonic sequences—and a relatively high degree of parametric similarity *within* each one—before the sequence as a whole can be heard polyphonically."

PRINCIPLE OF EQUIVALENCE. See EQUIVALENCE.

PROXIMITY-FACTOR. One of the primary gestalt-factors of cohesion and segregation described in Section II, and formulated there as follows: "In any collection of sounds (elements or clangs), those which are simultaneous or contiguous [in time] will tend to form perceptual groups (clangs or sequences), while relatively greater separations in time will produce segregation—other factors being equal."

REPETITION-FACTOR. One of the secondary factors of cohesion and segregation: "If a repetition of parametric profile is perceived within a series of sound-elements, this alone may produce a subdivision of the whole series into units corresponding to the repeated shape—the perceptual separation between the units occurring at the point just before the first repeated element."

RESONANT CLANG. A sort of borderline phenomenon—between the clang and the sequence—similar to the clang in many respects, but lasting so long that it functions as a (monomorphic) sequence, rather than as a real clang.

RHYTHM. See EXPLICIT and IMPLICIT RHYTHM.

RHYTHMIC INERTIA. A special form of the factor of objective set. It was said, in Section II, to involve "a psychological or kinesthetic tendency toward rhythmic repetition—the maintenance of a previously established rhythmic structure. . ." etc.

SCALE, PARAMETRIC. See PARAMETRIC SCALE.

SCALE, TEMPORAL. See PERCEPTUAL LEVEL.

SEQUENCE. Generally, "a succession of clangs which is set apart from other successions in some way, so that it has some degree of unity and singularity, constituting a musical gestalt on a larger perceptual level or temporal scale—though it will not be as 'strong' a gestalt as is the clang." This definition refers to the *polymorphic sequence*, (the *monomorphic sequence* being considered an exceptional case, not justifying the more generalized definition of *sequence* that would be necessary to include it). All sequences may be assumed to be comparable, however, with respect to *duration*—if only in that they tend to be longer than the clang, or longer than the normal range of durations within which it is possible to perceive an aural gestalt in one "grasp" of the attention. The gestalt-character of the sequence must therefore depend upon memory for its apprehension.

SEQUENCE-TYPES. See MORPHOLOGICAL RELATIONS.

SET. A psychological condition which may alter or modify the perception of a thing, as a result of previous experience. See OBJECTIVE SET and SUBJECTIVE SET.

SHAPE. An aspect of the form of a clang or sequence that is produced by the changes in parametric values from one moment to the next within the configuration. It has sometimes been used synonymously with such words as "profile," "contour," "outline," etc., even though there are obvious differences between the meanings of each of these terms in the realm of visual perception, from which they are borrowed. And none of them can mean quite the same thing there as they do in music—or as they are intended to mean in this book. But it is hoped that they will all connote approximately the same thing to the musician—that "aspect of form" referred to in the definition given above.

SIMILARITY-FACTOR. One of the primary gestalt-factors of cohesion and segregation described in Section II, and formulated there as follows: "In any collection of sound-elements (or clangs), those which are similar (with respect to values in some parameter) will tend to form clangs (or sequences), while relative dissimilarity will produce segregation—other factors being equal." The factor of similarity is probably the most important of all the gestalt-factors described, because (1) it applies to all parameters (the one in which this factor is manifested being called the *cohesive parameter*)—and even to higher-order "attributes" such as shape or form; (2) it is effective at many perceptual levels or temporal scales, from element and clang, to whole movements and pieces; and (3) it can function in both the horizontal (i.e., the temporal) and the vertical dimensions, and is the most effective factor in the differentiations necessary to any polyphonic texture.

STATISTICAL FEATURES. Overall, or "average" characteristics of a clang, such as parametric state, range (in each parameter), and duration of the clang as a whole—to be distinguished from the more specific, formal or morphological features of the clang.

SUBJECTIVE SET. Another of the secondary gestalt-factors—"expectations or anticipations [arising during a musical experience] which are the result of experiences *previous* to those occasioned by the particular piece of music now being considered."

TEMPORAL SCALE. See PERCEPTUAL LEVEL.

TIME-ENVELOPE. See ENVELOPE.

Bibliography

Artaud, Antonin. *The Theater and Its Double*. New York: Grove Press, 1958.

Cherry, Colin. *On Human Communication*. Cambridge: The MIT Press, 1978.

Ellis, Willis, D., ed. *A Source Book of Gestalt Psychology*. New York: Humanities Press, 1967. (Contains papers by Wertheimer and Köhler).

Joyce, James. *A Portrait of the Artist as a Young Man*. New York: The Viking Press, 1968.

Koffka, Kurt. *Principles of Gestalt Psychology*. London: Routledge & Kegan Paul Ltd., 1962.

Köhler, Wolfgang. *Introduction to Gestalt Psychology*. New York: New American Library, Mentor Books, 1959.

Schaeffer, Pierre. *A la Recherche d'une Musique Concrète*. Paris: Editions du Seuil, 1952.

Schoenberg, Arnold. *Style and Idea*. Leonard Stein, ed. London: Faber & Faber, 1975.

Seashore, Carl. *Psychology of Music*. New York: Dover Publications, Inc., 1967.

Thompson, D'Arcy W. *On Growth and Form*. Cambridge: University Press, 1968.

List of Musical Works Cited, and their Publishers

Bartók, Béla.
- *Sonate* (piano). Vienna: Universal-Edition, 1976.
- *4th String Quartet*. Vienna: Universal-Edition, 1939.

Ives, Charles.
- *Second Pianoforte Sonata ("Concord, Mass., 1840-60")*. New York: Associated Music Publishers, 1947.
- *Scherzo ("Over the Pavements")*, (chamber orchestra). New York: Peer International Corporation, 1954.
- *Three Places in New England, an Orchestral Set*. Bryn Mawr: Mercury Music Corporation, 1976.

Ruggles, Carl.
- *Evocations ("Four Chants for Piano")*. New York: American Music Edition, 1957.

Schoenberg, Arnold.
- *Drei Klavierstücke*, op. 11. Vienna: Universal-Edition, 1910.
- *Five Pieces for Orchestra*, op. 16. New York: Peters Corp., 1952.

Varèse, Edgard.
- *Density 21.5* (flute solo). New York: Colfranc Music Publ., 1966.
- *Octandre* (chamber ensemble). New York: Colfranc Music Publ., 1980.

Webern, Anton.
- *Sechs Stücke* (orchestra), op. 6. Vienna: Universal-Edition, 1961.
- *Fünf Stücke für Orchester*, op. 10. Vienna: Universal-Edition, 1951.
- *Fünf Sätze für Streichquartett*, op. 5. Vienna: Universal-Edition, 1949.

META Meta / Hodos

Preface

META Meta+Hodos represents an attempt to organize certain ideas first presented in *Meta+Hodos* in 1961, incorporating insights and revisions that have emerged since then. The writing was initially motivated by the desire to provide an outline of my ideas and terminology for use by students in a class in Formal Perception and Analysis at the California Institute of the Arts. The intent was therefore to make it as concise as possible, even if at the expense of comprehensibility, and I am aware that the result is probably not easily penetrated by someone not already familiar with *Meta+Hodos*. Nevertheless, I am pleased with the form it has taken, and hope that others may find it of interest in spite of its difficulties.

James Tenney
November, 1975

A. On Perceptual Organization

PROPOSITION I: In the process of musical perception, *temporal gestalt-units* (TG's) are formed, at several different hierarchical levels.

COMMENT I.1: The number of hierarchical levels in a given piece, and the relative durations of the TG's at adjacent hierarchical levels varies, depending on such things as style, texture, tempo, the duration of the piece, etc.

COMMENT I.2: TG's at a given hierarchical level are not always or necessarily disjunct—i.e., there are frequent intersections and ambiguities in their perceptual formation.

DEFINITION 1: A TG at the lowest (or first) hierarchical level will be called an *element*.

COMMENT 1.1: An element is a TG which is perceived as (temporally) *singular*—i.e., not divisible into lower-level (shorter) TG's. (See Comment IV.1.3, below, for a further description of element characteristics).

DEFINITION 2: A TG at the next higher (2nd) hierarchical level will be called a *clang*.

COMMENT 2.1: A clang is a TG at the lowest hierarchical level within which still-lower-level TG's are perceived.

DEFINITION 3: A TG at the next higher (3rd) hierarchical level will be called a *sequence*.

COMMENT 3.1: A clang thus consists of a temporal succession of two-or-more elements; a sequence consists of a temporal succession of two-or-more clangs. Note that a combination of two-or-more elements occurring simultaneously does not necessarily constitute a clang.

(For the case of simultaneous TG's see Definitions 5 through 8, below).

DEFINITION 4: The TG at the highest hierarchical level is the *piece* as-a-whole (but see Proposition V and Comment V.1, below).

COMMENT 4.1: The number of intermediate hierarchical levels (between those of the sequence and the piece) is variable (cf. Comment I.1, above).

DEFINITION 5: A TG whose component, next-lower-level TG's are perceived one-at-a-time will be called *monophonic*.

DEFINITION 6: A TG whose component, next-lower-level TG's are perceived two-or-more-at-a-time will be called *polyphonic*.

DEFINITION 7: A TG whose component TG's at all lower levels are monophonic will be called *simple*.

DEFINITION 8: A TG whose component TG's at any lower level are polyphonic will be called *compound*.

COMMENT 8.1: These terms will frequently be combined to describe four types of "vertical" construction or texture:

(1) a *simple-monophonic* TG (at a given hierarchical level) is one whose component TG's are monophonic (at all lower levels) and are perceived one-at-a-time (at the given level);
(2) a *simple-polyphonic* TG (at a given hierarchical level) is one whose component TG's are monophonic (at all lower levels) but perceived two-or-more-at-a-time (at the given level);
(3) a *compound-monophonic* TG (at a given hierarchical level) is one whose component TG's are polyphonic (at any lower level) but are perceived one-at-a-time (at the given level);
(4) a *compound-polyphonic* TG (at a given hierarchical level) is one whose component TG's are polyphonic (at any lower level) and are perceived two-or-more-at-a-time (at the given level).

COMMENT 8.2: The relationships among these four types of texture at three adjacent hierarchical levels are shown schematically in Figure 1.

PROPOSITION II: The perceptual formation of TG's at any hierarchical level is determined by a number of factors of cohesion and segregation, the most important of which are *proximity* and *similarity*; their effects may be described as follows:

PROPOSITION II.1: Relative temporal *proximity* of TG's at a given hierarchical level will tend to group them, perceptually, into a TG at the next higher level.

PROPOSITION II.2: Relative *similarities* of TG's at a given hierarchical level will tend to group them, perceptually, into a TG at the next higher level.

PROPOSITION II.3: Conversely, relative temporal *separation* and/or *differences* between TG's at a given hierarchical level will tend to segregate them into separate TG's at the next higher level.

COMMENT II.3.1: The perceptual formation of *lower-level* TG's is also affected by several secondary factors of cohesion and segregation, including *accent*, *repetition*, "*objective set*," and "*subjective set*" (see *Meta+Hodos*), but these will not be dealt with here.

B. On Musical Parameters.

DEFINITION 9: A *parameter* will be defined here as any distinctive attribute of sound in terms of which one sound may be perceived as different from another, or a sound may be perceived to change in time.

COMMENT 9.1: This definition refers to "subjective" or *musical* parameters (e.g., pitch, loudness, etc.) as distinct from "objective" or *acoustical* parameters (frequency, amplitude, etc.).

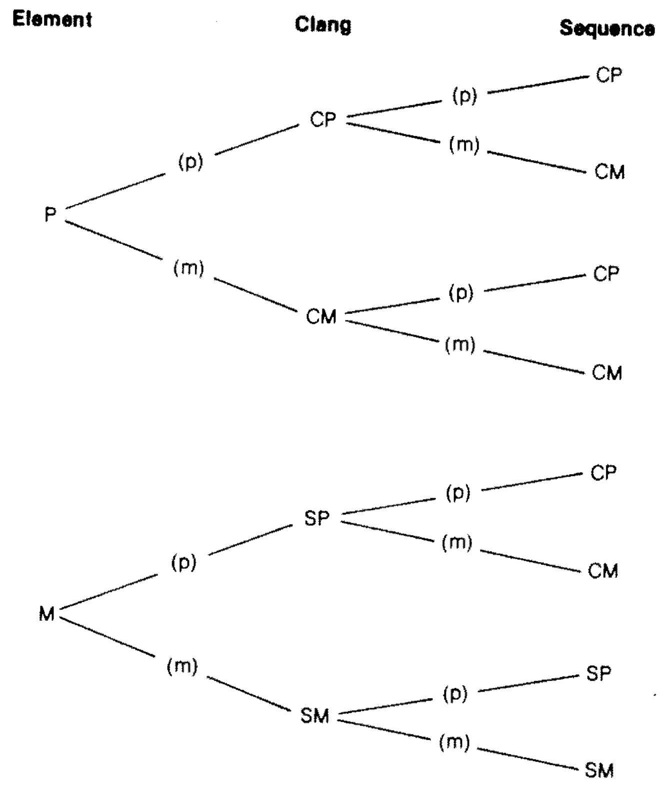


Figure 1. Relationships between simple, compound, monophonic, and polyphonic TG's at three HL's (M = monophonic, P = polyphonic, S = simple, C= compound, (m) = perceived one-at-a-time, (p) = perceived two-or-more-at-a-time).

COMMENT 9.2: There is not, in general, a one-to-one correspondence between musical and acoustical parameters. Where there *is* such a correspondence, the relation is more nearly logarithmic than linear.

PROPOSITION III: Pitch, timbre, and (musical) time are not simply one-dimensional parameters, because each includes at least two relatively independent "sub-parameters."

COMMENT III.1: Similarities and differences between any two pitch *intervals* are perceived in two different ways, depending on their relative magnitudes and their interval qualities. These, in turn, result from differences in what will be called (1) *pitch-height*, and (2) *pitch-chroma*.

DEFINITION 10: *Pitch-height* refers to that aspect of pitch-perception which depends on the existence of a continuous range of pitches, from low to high.

DEFINITION 11: *Pitch-chroma* refers to that aspect of pitch-perception which depends on the phenomenon of "octave equivalence," and the fact that the continuous range of pitches is also *cyclic*, virtually returning to its starting-point in each transition from one octave to the next.

COMMENT 11.1: These two sub-parameters may be related to the fact that there are two distinct mechanisms of pitch-perception involved in hearing—a "place" mechanism (determining pitch-height) and a "time" mechanism (determining pitch-chroma). The place mechanism is most effective for high frequencies, the time mechanism for lower ones, but the two overlap over a fairly broad range in the middle register, and it is here that our pitch-perception is the most acute (and the most bi-dimensional).

COMMENT 11.2: The multi-dimensionality of *timbre* is due to the fact that it is determined in a complex way by our perception of a large number of acoustical features, which may be subsumed under three categories:

(1) the steady-state *spectrum*,
(2) various kinds of steady-state *modulations*, and
(3) transient modulations or *envelopes*

COMMENT 11.3: The sub-parameters of (musical) *time* will be called (1) *epoch*, (2) *duration*, and (3) *temporal density*.

DEFINITION 12: *Epoch* refers to the moment of occurrence—in the ongoing flow of experienced time—of any musical "event," compared to some reference moment such as the beginning of the piece.

DEFINITION 13: The *temporal density* of a TG is the number of its component, next-lower-level TG's per unit time; ("*duration*" will be used in its usual sense).

COMMENT 13.1: The average temporal density of a TG at a given hierarchical level will thus be equal to the reciprocal of the average duration of its component TG's at the next lower level.

COMMENT 13.2: "Tempo" is a special case of temporal density, referring to an expressed or implied pulse or "beat," rather than to actual durations, and it is only relevant to *lower-level* TG's.

DEFINITION 14: Pitch-height and epoch (which correspond most closely to the acoustical parameters, log-frequency and "real" time) will be called *distributive* parameters, because a difference in at least one of these is necessary for two sounds to be perceived as *separate*.

DEFINITION 15: All other parameters (including loudness, pitch-chroma, duration, temporal density, and the several sub-parameters of timbre) will be called *attributive* parameters. Note that a difference in any of these is insufficient, by itself, for two sounds to be perceived as separate—there must *also* be a difference in one of the distributive parameters.

C. On Formal Perception and Description

PROPOSITION IV: The perception of form at any hierarchical level involves the apprehension of three distinct aspects of form, at that and all lower levels. These three aspects of form will be called *state*, *shape*, and *structure*.

DEFINITION 16: *State* refers to the statistical and other "global" properties of a TG, including the *mean values* and *ranges* in each parameter, and its *duration*.

DEFINITION 17: *Shape* refers to the "profile" of a TG in some parameter, determined by *changes* in that parameter with respect to either of the distributive parameters, epoch and pitch-height (or their acoustical correlates, "real" time and log-frequency).

DEFINITION 18: *Structure* refers to *relations between subordinate parts* of a TG—i.e., relations between its component TG's at the next (or several) lower level(s). (See also Definition 19 and its Comments, below.)

PROPOSITION IV.1: A complete description of a monophonic TG at any hierarchical level requires descriptions of state, shape, and structure, for every parameter with respect to *time*.

COMMENT IV.1.1: In this context (i.e., that of monophonic TG's), shape is time-dependent, while state and structure are "out-of-time" characteristics (but see Comment IV.2.1, below).

COMMENT IV.1.2: The state of a monophonic TG simply depends on lower-level *states*; shape is determined by *changes of state* at the next lower level; structure depends on *relations between states, shapes and structures* at the next (or several) lower level(s) (see Figure 2).

COMMENT IV.1.3: Since, by Definition 1, Comment 1.1, an element is not perceived as "divisible into lower-level TG's," the structure of an element is not

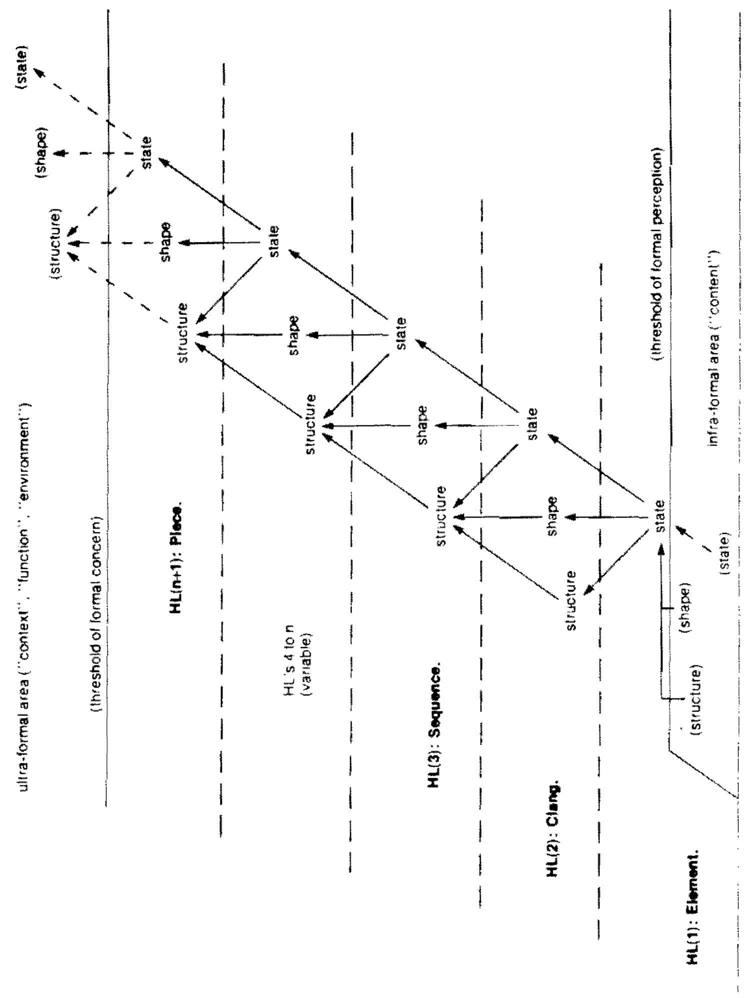


Figure 2. Relationships between the three aspects of form at several hierarchical levels (HL's).

perceived directly—i.e., element-"structure" is located in the "infra-formal" area of Figure 2, below the "threshold of formal perception." Element-"shape" is sometimes above, sometimes below this threshold.

COMMENT IV.1.4: The various state-descriptions of an element are equivalent to the set of parametric values needed to describe the element (except when aspects of element-shape are also reduced to parameters—e.g., amplitude-envelope shape).

COMMENT IV.1.5: The "similarities" and "differences" of Propositions II.2 and II.3 may be of all three kinds—i.e., of state, shape, or structure.

DEFINITION 19: There are three basic types of structure (corresponding to the three connecting lines to "structure" in Figure 2). These will be called

(1) *statistical structure* (i.e., relations between lower-level states),
(2) *morphological structure* (relations between lower-level shapes), and
(3) *cascaded structure* (relations between lower-level structures).

COMMENT 19.1: Each of these three types of structure may be specified by showing the relations between each lower-level TG and *every other TG* at that level. For a given set of relations (limited in such a way that there is only one relation between each pair of TG's), this might be done by arranging them in a square array or matrix. In the case of statistical structure, such a matrix might show, for example, the set of *intervals* between the parametric mean values of each pair of TG's.

COMMENT 19.2: For morphological structure, the relations included in such a matrix might be as few as three (e.g., =, ≠ and T, for "identical to," "unrelated to," and "related via some transformation," respectively), or the "T" might be expanded into a longer list such as the following:

E/C (for expansion/contraction of intervals),
X/L (extension/elision at the ends of a TG),
I/D (interpolation/deletion into or from within a TG),
I (inversion),
R (retrogression),
W ("warping" or distortion of shape, still preserving its essential topological features),
P (permutation of the order of component TG's), etc.

COMMENT 19.3: For cascaded structure, the only relations needed for such a matrix might be = and ≠.

DEFINITION 20: In addition to the three basic types of structure listed in Definition 19, there is still another type which is relevant to musical perception, one involving relations between *relations*, rather than relations between (various aspects of) the TG's themselves. These will be called *relational structures*, and may be of three kinds: (1) *state-relational* structure, (2) *shape-relational* structure, and (3) *structure-relational* structure.

PROPOSITION IV.2: A complete description of a polyphonic (or compound-monophonic) TG at any hierarchical level requires descriptions (in addition to those listed in Proposition IV.1) of state, shape, and structure for each of the *attributive* parameters with respect to *log-frequency*.

COMMENT IV.2.1: In this context, although shape is not time-dependent, it still involves the sequential order of states in the frequency domain; state and structure do not.

COMMENT IV.2.2: For polyphonic TG's, the relationships between state, shape, and structure (with respect to frequency)—such as those described in Comment IV.1.2, above—are not yet known.

PROPOSITION V: Formal properties at a given hierarchical level determine the (non-semantic) "content" of the TG's at the next higher level; they also determine the "context" (or "function" or "environment") of TG's at the next lower level.

COMMENT V.1: What we do finally *call* (non-semantic) "content" is the result of "forms" at a level below the first one we are able to perceive "formally"; what we call "context" (or "function" or "environment") is determined by formal conditions at a level above the largest one we *choose to deal with* "formally."

PROPOSITION VI: As we move from the infra-formal area up into and through the first few specifically formal hierarchical levels, new parameters emerge.

COMMENT VI.1: Even *within* the infra-formal area there is a similar "emergence"—e.g., the transition from the basic physical nature of the signal as (simply) amplitude vs. time, to the (acoustical) parameter, *frequency*. Examples above the threshold include the *timbre*-effects of rise-time and vibrato (at HL(1) in Figure 2), and *temporal density* at HL(2)).

PROPOSITION VII: There is a close correlation between what may be called *parametric focus* and the relative *range of variation* of next-lower-level states within a TG, i.e., the greater the range in a given parameter, the more one's attention will be focussed on the changes in that parameter, and the more prominent will be the shape determined by those changes.

DEFINITION 21: A parameter whose variation (over a relatively wide range) at the next lower level thus focusses the attention on the shape of a TG in that parameter will be called a *formative* parameter.

DEFINITION 22: A parameter whose relative constancy (or variation over a *narrow* range) at the next lower level is thus significantly responsible for its unity as a gestalt (via the similarity-factor of Proposition II.2) will be called a *cohesive* parameter.

PROPOSITION VIII: The formative parameters of a TG are generally different from the cohesive parameters of that same TG.

COMMENT VIII.1: This follows almost simply "by definition," but its implications are important enough to justify it as a separate Proposition.

PROPOSITION IX: The formative parameters of a TG at a given hierarchical level are generally different from the formative parameters of the next-higher-level TG which contains it.

COMMENT IX.1: One obvious exception to Propositions VIII and IX may occur when the formative parameter of a TG is pitch, but this is only possible because the number of distinguishable values in this parameter is very great—and it can only occur when the range of pitch-variation within the next-lower-lever TG's is relatively limited. The more extensive the range covered within each lower-level TG, the less perceptible will be the changes of pitch-state from one TG to the next, and thus the less effective will pitch be as a formative parameter at the next higher level. This adjacent-level "trade-off" relation is made more explicit and precise in the following Proposition:

PROPOSITION X: For any parameter with respect to time, the greater the range of variation at a given hierarchical level, the smaller the range of variation possible at the next higher level, and vice versa.

COMMENT X.1: For a given parameter, and under the special condition that the ranges are identical for all TG's at a given hierarchical level, the following relations will hold:

For the first hierarchical level, considered by itself, the maximum range available is $N(1)_{max} = N_t$, where N_t is the total number of distinguishable values in that parameter. When two hierarchical levels are considered, the maximum range at the second level is

$$N(2)_{max} = N_t - (N(1) - 1).$$

For a third level, the maximum range will be

$$N(3)_{max} = N_t - (N(1) - 1) - (N(2) - 1).$$

More generally, the maximum range available at a given level (L) is

$$N(L)_{max} = N_t - (N(1) - 1) - (N(2) - 1) - \ldots (N(L - 1) - 1), \text{ or } N(L)_{max} = N_t - N_L + L - 1.$$

Finally, the total available range (N_t) may be distributed equally among some number of levels (L), so that

$$N(1) = N(2) = \ldots N(L), \text{ and } N(L+1)_{max} = 0,$$
by setting each N at $N = N_t/L+1$.

DEFINITION 23: A TG whose component, next-lower-level TG's all have the *same state* in a given parameter will be called *ergodic* with respect to that parameter.

COMMENT 23.1: The shape of an ergodic TG is thus "flat" in that parameter.

COMMENT 23.2: An ergodic TG has the same parametric state as each of its component, next-lower-level TG's.

DEFINITION 24: A TG whose component, next-lower-level TG's have *different states* in a given parameter will be called *non-ergodic* with respect to that parameter.

COMMENT 24.1: The *shape* of a TG may thus be either ergodic or non-ergodic, with respect to a given parameter.

DEFINITION 25: A TG whose component, next-lower-level TG's all have the *same shape* in a given parameter will be called *isomorphic* with respect to that parameter.

DEFINITION 26: A TG whose component, next-lower-level TG's all have *different* (or more precisely, *unrelated*) *shapes* in a given parameter will be called *heteromorphic* with respect to that parameter.

DEFINITION 27: A TG whose component, next-lower-level TG's have shapes that are related to each other via some process of *transformation* will be called *metamorphic* with respect to that parameter.

COMMENT 27.1: The *morphological structure* of a TG may thus be either isomorphic, heteromorphic, or metamorphic, with respect to a given parameter.

D. On Entropy As A Measure Of Variation.

DEFINITION 28: One of the most important aspects of musical experience is the perception of variation, and a useful measure of variation is *entropy*. In information theory, the entropy of a "message" consisting of a series of n discrete "symbols" drawn from an "alphabet" of N equally probable symbols is

$$H = n \log_2 N \text{ (bits per message).}$$

The entropy of each symbol is

$$H = \log_2 N \text{ (bits per symbol).}$$

COMMENT 28.1: The most important variable here is N, the number of symbols available. In the special case where $N = 1$, $H = 0$.

COMMENT 28.2: When the available symbols are not equally probable—i.e., when they do not occur with the same relative frequencies (p_i)—then

$$H = -\Sigma\, p_i \log_2 p_i \text{ (bits per message).}$$

DEFINITION 29: We may define as many different types of entropy as there are different types of structure. Thus, we may distinguish between *statistical, morphological*, and *structural entropies*, according to whether the "symbols" considered are lower-level *states, shapes*, or *structures*. In addition, there will be three *relational entropies*—those involving state-relations, shape-relations, and structure-relations.

DEFINITION 30: The entropies of a TG at a given hierarchical level may be measured either in terms of component TG's at the lowest (i.e., element-) level, or in terms of component TG's at the next lower level. The first kind of measure (which has been the usual procedure in most applications of information theory) will be called an *additive* measure, the second (which will be used most often here) will be called an *adjacent-level* measure of entropy.

DEFINITION 31: Since a TG at every hierarchical level except the lowest and highest (i.e., any except an element or the whole piece) may be considered *both* a message (containing lower-level symbols) *and* a symbol (contained within a higher-level message), the various entropies may be defined for a TG either as *message-entropies* or as *symbol-entropies*.

COMMENT 31.1: The following Propositions refer to adjacent-level message-entropies of a TG:

PROPOSITION XI: The statistical entropy of an ergodic TG is zero.

PROPOSITION XII: The state-relational entropy of an ergodic TG is zero.

PROPOSITION XIII: The statistical entropy of a non-ergodic TG at a given hierarchical level depends on
(1) the number of its component, next-lower-level TG's.
(2) the number of their distinguishable states, and
(3) the relative frequencies of these states.

PROPOSITION XIV: The state-relational entropy of a non-ergodic TG at a given hierarchical level depends on
(1) the number of its component, next-lower-level TG's.
(2) the number of the distinguishable differences between their states, and
(3) the relative frequencies of these differences.

PROPOSITION XV: The maximum statistical entropy attainable in a TG at a given hierarchical level is *inversely related* to the statistical entropy of its component TG's at the next lower level. (This is a consequence of Proposition X.)

PROPOSITION XVI: The morphological entropy of an isomorphic TG is zero.

PROPOSITION XVII: The shape-relational entropy of an isomorphic TG is zero.

PROPOSITION XVIII: The morphological entropy of a heteromorphic TG is maximal (for a given number of next-lower-level TG's).

PROPOSITION XIX: The shape-relational entropy of a heteromorphic TG is zero.

> COMMENT XIX.1: There must be a meaningful way to define the morphological entropy of a metamorphic TG, but this has not yet been found.
>
> COMMENT XIX.2: Nothing is yet known about structural entropies.