Music Perception
Spring 1996, Vol. 13, No. 3, 401–432

A Perceptual Analysis Of Mozart's Piano Sonata K. 282: Segmentation, Tension, and Musical Ideas

CAROL L. KRUMHANSL
Cornell University

The experiments reported here provide a perceptual analysis of the first movement of Mozart's Piano Sonata in E♭ Major, K. 282. The listeners, who varied in the extent of their musical training, performed three tasks while listening to the piece as it was reproduced from an expert performance. The first task determined how the music is perceived to be segmented, the second task determined how the experience of tension varies over time, and the third task determined what listeners identify as new musical ideas in the piece. These tasks were performed first on the entire piece and then on smaller sections from the beginning. These three aspects of music perception are coordinated with one another and correlate with various musical attributes. Judgments of section ends co-occurred with peaks in tension and slow tempos. Judgments of new musical ideas co-occurred with low tension levels and neutral tempos. Tension was influenced by melodic contour, note density, dynamics, harmony, tonality, and other factors. Judgments of large-scale section ends were less frequent than judgments of new musical ideas, but these were more nearly one-to-one on smaller time scales. A subsidiary experiment examined the extent to which tension judgments were influenced by performed tempo and dynamics. Listeners made tension judgments for four different versions of the piece: as performed, constant dynamics (with tempo as performed), constant tempo (with dynamics as performed), and constant tempo and dynamics. The tension curves were generally very similar, deviating only in a few regions containing major section ends. The results are considered in light of the metaphor of tension applied to music and the analogy between music and linguistic discourse.

THE experience of music is notoriously difficult to describe. As a consequence, a wide variety of different approaches to musical analysis have been developed (as summarized, for example, by Bent, 1987, 1994; Cook, 1987). Each approach has its guiding metaphors, special terminology, descriptive devices, and theoretical commitments. Some are oriented toward specific musical styles or compositional methods and are narrowly focused on musical concerns. Others engage broader philosophical and psychologi-

Requests for reprints may be sent to Carol L. Krumhansl, Department of Psychology, Uris Hall, Cornell University, Ithaca, NY 14853. (e-mail: clk4@cornell.edu)

cal questions. In any case, the primary objective of musical analysis is the explication of particular musical compositions, and a method's success is measured by whether it yields useful insights when applied to those compositions. This article explores the possibility that perceptual experiments might serve as a useful complement to other approaches to the analysis of music.

In the experiments reported here, listeners performed a number of perceptual tasks while they listened to the first movement of Mozart's Piano Sonata in E♭ Major, K. 282. The tasks were designed to probe three aspects of music perception: segmentation into hierarchically organized units, variations over time in the degree of experienced tension, and identification of distinct musical ideas as they are introduced in the piece. The tasks were designed so that they did not require the listeners to have musical expertise. The listeners who participated varied considerably in their musical backgrounds, but few had explicit training in musical analysis. Thus, the experimental results are independent of results obtained through other methods. Because the experiments investigated the perception of this single piece of music, the findings do not extend to other compositions or styles, although the tasks would seem applicable to a range of styles. The broader context for the work draws on two different traditions. One is the description of music as giving rise to variations in tension over time. The other is the analogy between music and discourse.

Musical Tension

Music, particularly music in the Western tonal-harmonic style, is often described in terms of patterns of tension and relaxation (or release from tension). According to Bent (1987), this way of describing music arose in early 20th century musical analysis and was influenced by Gestalt psychology. Three Gestalt principles were extended to music: closure, which automatically completes partially incomplete patterns; the phi phenomenon, which interpolates a link between two separate occurrences; and *Prägnanz*, which seeks the simplest possible perceptual organization. In addition, figure-ground organization motivated the analysis of music by using reductions to show the essential underlying structure. Musical form was considered a type of whole or Gestalt. Phrases, motives, rhythms, and other musical patterns bear specific relationships with one another, and together determine the musical form. The musical form, in turn, exerts influences on the perception of the musical components, just as a visual form influences the perception of its components. Finally, this tradition of music analysis is congenial with Gestalt theory in its reliance on graphical methods.

One of the leading proponents of this approach was Ernst Kurth. In his terminology, melodic, harmonic, and rhythmic elements carry energy, which

1991). Graphical expressions of dynamic qualities in musical responses can be found in the work of Truslit (see Repp, 1993) and Clynes and Nettheim (1982).

The most extensive experimental analysis of musical tension, to my knowledge, however, is the dissertation by Nielsen (1983), which has recently been extended in studies by Madson and Fredrickson (1993; Fredrickson, 1995). According to Nielsen, "In the musical structure of strata, 'tension' is assumed to be placed in the middle region of the object, connecting structural characteristics of the surface level with more deeply located strata of emotion and other strata of meaning" (p. 316). In the experiments, listeners pressed a pair of tongs together to indicate the experienced degree of tension. The amount of pressure applied was recorded continuously as the music was played. The musical examples used were the first movement of Haydn's Symphony, no. 104, and the first 75 measures of R. Strauss' *Also sprach Zarathustra*. Listeners were experienced musicians and 16-year-old students. This method produced strikingly regular tension curves, with smaller waves of tension superimposed on larger waves of tension. Intersubject agreement was reasonably strong. Greater consistency was found among the experienced musicians, although some of the students' tension curves matched those of the musicians. The tension judgments could be related to specific musical factors, including dynamics, timbre, melodic contour, harmony, tonality, and repetition. These factors, Nielsen acknowledges, interact in complex ways and the tension that is felt is assumed to result from the interactions.

Listeners in the experiments to be reported here performed a similar task. As they listened to the movement from the Mozart sonata, they adjusted an indicator on the computer display to show the degree of experienced tension. The position of the indicator was recorded four times per second. The computer control of the music permitted precise registration between the musical events and the responses. This task was repeated a number of times, first with the entire piece, then with smaller subsections. This gives a way of assessing the reliability of the tension judgments across repetitions as the listeners became more familiar with the piece. Listeners had various levels of musical training, so comparisons across individuals examine differences that depend on prior musical experience.

Music and Discourse

Descriptions of music have often relied on analogies between music and language. According to Bent (1994), the earliest surviving full-scale analysis of music, from the Middle Ages, identifies rhetorical sections and other quasi-linguistic units. More recent descriptions have drawn on various parts

This description shares much with accounts of musical organization. In fact, Chafe says, "Once one has become accustomed to observing intonation units,...it becomes impossible not to hear analogous segments in music. Their presence there may be no accident. The convergence of language and music in this respect may very well show a human need to process information in relatively brief units in active consciousness, to combine such units within larger centers of interest, and every so often to shift from one cluster of semi-active information to another" (p. 186). As a demonstration of how his analytic techniques might be applied to music, Chafe presents brief analyses of the beginning of Mozart's Piano Sonata in F Major, K. 322, and a Seneca Indian song from *Drum Dance*. The intonational units identified in the analysis appear to be defined primary by melodic contour and pauses.

Two tasks were included in the experiments reported here to examine aspects of music that may be analogous to discourse. The first aspect is segmentation into hierarchically organized units. One task required listeners to indicate when they heard the end of sections within the piece, similar to the method used by Imberty (1981), Clarke and Krumhansl (1990), and Deliège and El Ahmadi (1990). This was done first for the entire piece, and then for smaller subsections. Of interest are the kinds of features that occur at the ends of sections and whether these changed with the focus on smaller subsections of the music. The second aspect is the introduction and subsequent elaboration of new materials or topics. A second task required listeners to indicate when they heard new musical ideas introduced in the music. Again, this was done first for the entire piece and then for smaller subsections. The responses can be used to determine musical features that set off new musical ideas and can be compared with music analytical concepts such as figure, motif, and phrase (see Bent, 1987). Finally, the positions of the new musical ideas in relation to section endings can be found, as well as how both of these relate to the tension judgments.

Experiments 1–3

The first three experiments all followed the same general method, and the same listeners participated in all three. In the first, listeners heard the entire piece and, on successive hearings, made three different kinds of judgments: section ends, tension, and new musical ideas. Listeners made the same judgments for measures 1–15 in Experiment 2 and measures 1–8 in Experiment 3. The three experiments were always run in the same order, so that experience with the piece increased throughout the series of experiments. The focus in the later experiments on shorter sections might also change the listeners' thresholds for identifying section ends and new musi-

cal ideas and for
experiments allo
different time sca

Subjects

Fifteen members
experiment. Their n
Radio Music Source
ably in their musical
musical instruction.
tion summed over a
struction was 12. Fiv
level. One listener re
piece but had not pla
the experiment.

Apparatus

The music was pl
software. The MIDI
sizer set to the Acou
was amplified by a Y
AKG headphones.

Stimulus Material

The stimulus mat
The recording on th
and then reproduced
dynamics accurately.
used to code the MI
the synthesizer uses
namics of the acoust
periment 1 used the
measures). Experime
the first beat of meas
8 measures of the pie

Procedure

The display on th
listeners heard the e
responses. In Step 2,
button in the center
piece. In Step 3, they
the amount of tensio
adjust it as frequentl
button in the center o
music.

Experiments 2 an
that they would perf

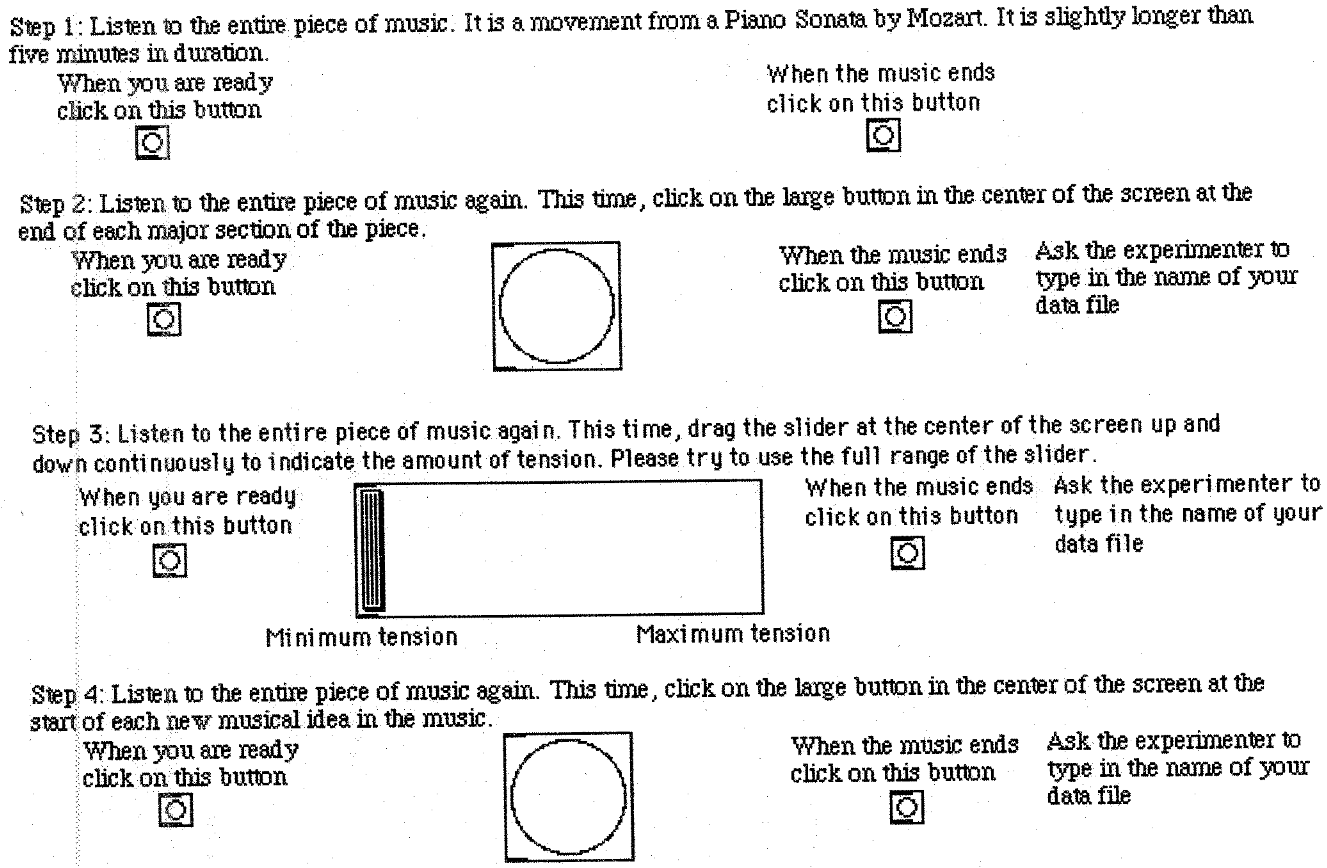

Fig. 1. The computer interface with which listeners made their responses in Experiment 1. After hearing the music (Step 1), they judged when large-scale section ends occurred (Step 2), the degree of tension (Step 3), and when new musical ideas occurred (Step 4). Experiments 2 and 3 adapted the instructions to smaller sections of the piece.

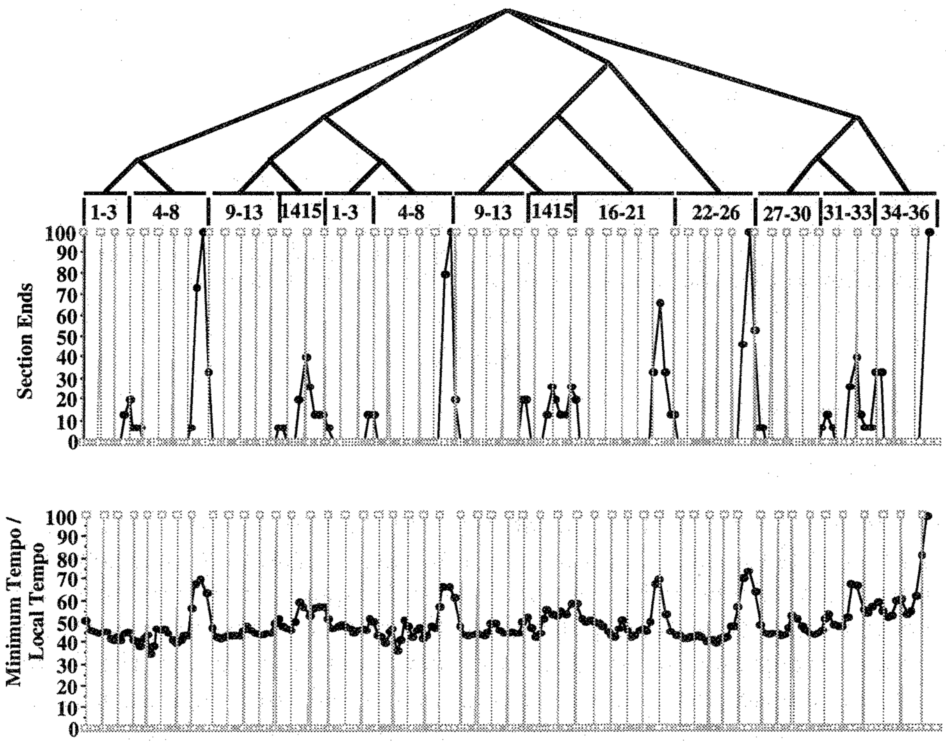

Fig. 2. The top graph shows the judgments of large-scale section ends in Experiment 1. The values shown are the percentages of listeners responding within each two-beat interval of time; the gray lines mark the measures. The tree structure shown at the top is derived from the section-end judgments. The bottom graph shows the duration of each two-beat interval divided by the longest two-beat interval in the piece. This is equivalent to the minimum tempo divided by the local tempo (which is expressed as a percentage). Higher values correspond to slower tempos. Judgments of section ends co-occur with slowing of tempo.

The perceptual judgments of segmentation correspond closely to the traditional analysis of form as described by Narmour (this issue, Figure 1). However, the tree derived from the perceptual judgments corresponds less well to the global and prolongational analysis by Lerdahl (this issue, Figure 17). Lerdahl assumes, in line with conventional wisdom, that the segmentation of measures 1–15 should be the same as measures 1–15 (repeat). However, because of the strong ending at the end of measure 8 (repeat), the perceptual judgments produce a tree in which measures 1–8 (repeat) join with the preceding measures 9–15, and measures 9–15 (repeat) join with the subsequent measures 16–26. It should be noted that the perceptual judgments do not readily suggest a way to derive the "tensing" versus "relaxing" relationship that is represented in the prolongational analysis.

The correspondence between large-scale segmentation and performed tempo was striking. Local tempo was measured over the same two-beat

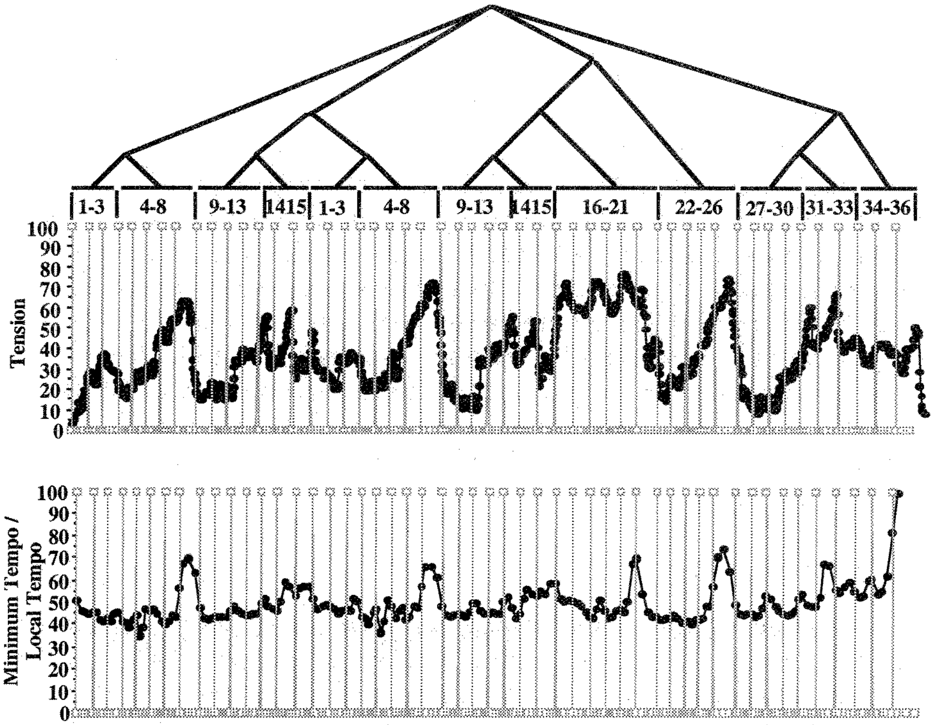

Fig. 3. The top graph shows the judgments of tension in Experiment 1. Listeners indicated tension by adjusting an indicator whose position was measured every 250 ms on a scale from 0 to 100. Comparison with the tree structure at the top shows tension peaks followed by rapid decreases at the ends of large-scale sections. Comparison with the bottom graph shows slower tempos at the same points in the music.

29, 31, and 33. A subsidiary experiment, described later, examines the effect of manipulating tempo on tension judgments.

Figure 4 shows the tension data in more detail. The top and middle graphs show the data for measures 1–15 and 1–15 (repeat), emphasizing the degree of consistency in the responses across repetitions. The material in measures 22–33 is analogous to that in measures 4–15 and, again, the degree of consistency is strong despite the surface differences between these sections. At a general level, the tension judgments correspond to the traditional account presented by Narmour (this issue) of a stable primary theme in measures 1–3 and a stable secondary theme in measures 9–11. However, he notes an alternative interpretation of measures 1–3 as introductory material and measures 4–6 as the primary theme. This is consistent with the relative lack of tension in measures 4–6.

Narmour (personal communication, May 1994) also provided a more detailed account of locations predicted to be low in tension and locations

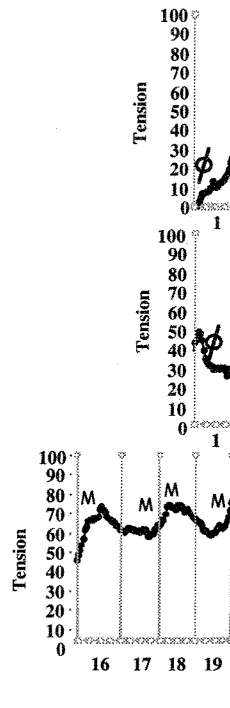

Fig. 4. Shows the ten
sistency in the tensio
in measures 4–15 and
munication, May 19
tural characteristics

and sources of te
table, the source
more sources wo
dicted to be low
sion or are soon f
tension, indicated
soon followed by
lar tension curve
tension are not al
and 24, and 7 an
confirmed by the
Comparing the
sion curves migh
the tension curve
with a schematic

TABLE 1
Sources of Tension in Narmour's Analysis

Location	Source of Tension
Measure 2	Dissonance
Measure 3	Extraopus harmonic style
Measure 4	Melody, dynamic
Measure 5	Melody, dynamic
Measure 6	Mode
Measure 7	Dissonance, dynamic
Measure 8	Melody
Measure 11	Denial of intraopus style
Measure 13	Denial of intraopus style, denial of extraopus harmonic style
Measure 14	Chromaticism, denial of intraopus style, denial of extraopus harmonic style
Measure 15	Melody
Measure 16	Denial of intraopus style, key change
Measure 17	Chromaticism, dynamic
Measure 18	Denial of intraopus style
Measure 19	Harmonic process, dynamic, chromaticism
Measure 20	Denial of intraopus style
Measure 21	Dissonance, denial of intraopus style
Measure 22	Denial of intraopus style
Measure 23	Break in pattern of harmony
Measure 24	Dissonance, dynamic
Measure 25	Melody, denial of intraopus style, dynamic
Measure 26	Melody
Measure 29	Denial of intraopus style
Measure 31	Denial of intraopus style, denial of extraopus harmonic style
Measure 32	Chromaticism, denial of intraopus style, denial of extraopus harmonic style
Measures 34, 35	Melody, dissonance

covary with the pitch height of the melody at a local level (the highest notes in the schematic). This was particularly so in measures 9–15, 16–19, and 27–33, where the fine-grained detail of the tension curves quite closely followed the melodic contour. This figure also makes clear that the sharpest drops in tension occurred when the density of notes decreases in measures 8, 21, 26, and 36. Figure 5 also displays the MIDI-coded key velocities under the tension graphs. As noted earlier, these values only approximate the dynamics of either the original performance or the version that was reproduced in the experiment. Even so, some correspondence was found. The fine-grained detail in tension in measures 9–15 and 27–33 tended to follow the highest velocity values. Also, the major section ends in measures 8, 15, 21, 26, and 36 were accompanied by declines in dynamics. Thus, influences of both dynamics and pitch height can be found in the tension profiles. It is interesting to note the correspondence between dynamics and

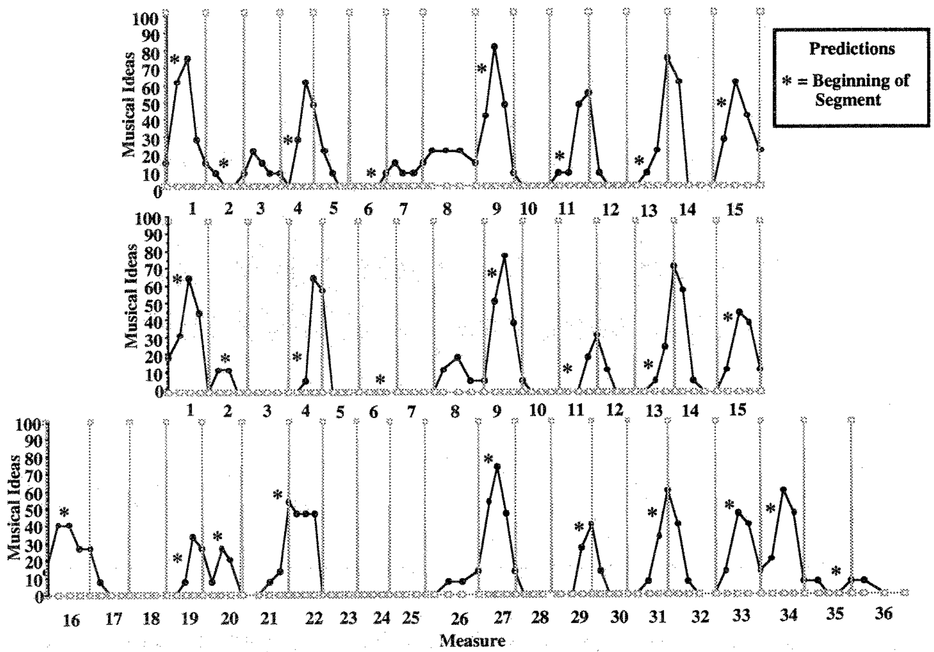

Fig. 6. Judgments of new musical ideas in Experiment 1. The values shown are the percentage of listeners responding within each two-beat interval of time. Listeners identified many new musical ideas within the piece. Comparison with Figure 2 shows that judgments of large-scale section ends were followed by judgments of new musical ideas. New musical ideas also tended to occur at neutral tempos. The three graphs in this figure indicate the consistency between repetitions and between the analogous material in measures 4–15 and 22–33. Theoretical predictions from Robert Gjerdingen (personal communication, March 1995) show the beginning of musical figures indicated by an asterisk.

age of listeners indicating that a musical idea began within each two-beat interval. Listeners apparently found a large number of different musical ideas in the piece. In most cases, there was reasonable consensus about their locations; the majority of peaks in the curve indicate agreement by at least half of the listeners.

Considering first the relationship with segmentation, section ends were in all cases followed by new ideas. Judgments of new ideas occurred at the beginning of the piece and in measures 4, 9, 14, 1 (repeat), 4 (repeat), 9 (repeat), 14 (repeat), 16, 22, 27, 32, and 34, that is, after each point at which section-end judgments occurred. However, new musical ideas also occurred within these segments, most notably in measures 11, 15, 11 (repeat), 15 (repeat), 29, and 33. Because of these additional judgments and the temporal lag between new musical idea and section-end judgments, the correlation between responses of section ends and musical ideas was nonsignificant [$r = -.06$ ($N = 202$, $df = 200$)]. As far as tempo is concerned, new musical ideas tended to occur at a neutral tempo. These variables had a

correlation of –
esting relations
can be seen by
occur at positio
markedly. This
peat), 4 (repeat
27, 29, 32, 33,
that occurred in
local decrease i

Figure 6 also
nication, March
musical figures.
an asterisk. Wit
were followed b
good agreemen
ments of new m
figural strategy
torical preceder
figures: Openin
middle of meas
14).

Comparison
sponded to mul
in rhythm and a
almost immedia
changed to an
times what liste
shift in the regi
ideas in various
changing the de
musical materia
tiple surface cha

Finally, when
of new musical
commented that
indicate a chang
the piece.

Experiments

Medium- and

Experiment 2
asked to judge
fore, the respon

ages of listeners who responded within each interval are shown at the center of Figure 7. The data from Experiment 1 are shown at the bottom of the figure for comparison. As before, listeners agreed unanimously that a section ending occurred in measure 8. The positions eliciting fewer responses in the previous experiment, in measures 3, 13, and the beginning of measure 15, now received more responses, and additional responses occurred in measures 2, 6, 7, and 11. Despite the additional responses, the correlation between Experiments 1 and 2 was high [$r = .81$ ($N = 58$, $df = 56$), $p < .0001$].

In Experiment 3, listeners heard only the first eight measures and judged where endings of small-sized sections occurred. The data are shown at the top of Figure 7. The same kinds of changes in responding were found as had been observed between Experiments 1 and 2. Measures 2, 6, and 7,

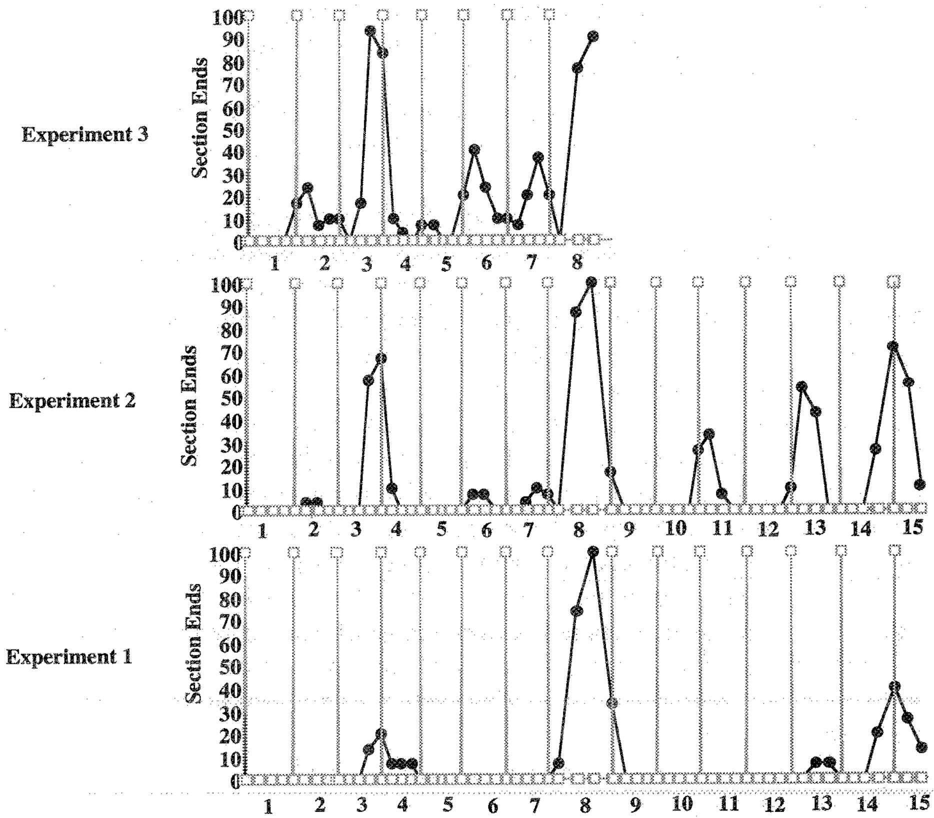

Fig. 7. Judgments of section ends made on the large scale (whole piece, Experiment 1), medium scale (measures 1–15, Experiment 2), and small scale (measures 1–8, Experiment 3). Focusing on smaller sections produced responses in more locations but the timing of the responses was approximately the same in all three experiments.

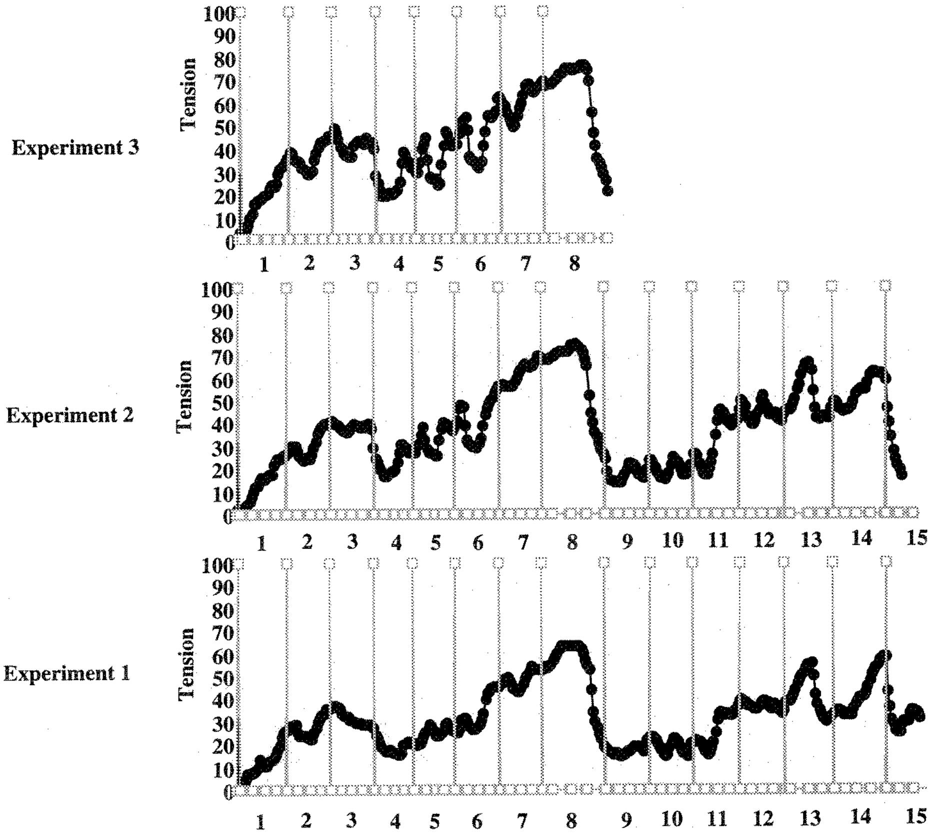

Fig. 8. Judgments of tension in Experiments 1, 2, and 3. Although the later experiments generated profiles with slightly more fine-grained structures, the patterns were strongly correlated.

than before, as can be seen by comparison with Figures 3 and 5. In addition, tension on the medium- and small scales was even more tightly coupled with judgments of section ends than in Experiment 1. Peaks in tension coincided with section-end judgments in measures 2, 3, 6, 7, 8, 13, and 15 in Experiment 2, and additionally in measure 5 in Experiment 3. The only disparity was in measure 11.

The tension ratings for the first eight measures were averaged over the three experiments and are shown in Figure 9. For this opening segment, Lerdahl (this issue, Figure 21) provided a set of numerical predictions for tonal tension of each event. The definitions of the terms and the rationale for their numerical coding appear in his article. Briefly, *scale degree* codes whether the melodic tone is contained in the supporting triad. *Inversion* codes whether the triad is in root position or inversion. *Nonharmonic tone* codes for the presence of tones not in the chord, including chordal sev-

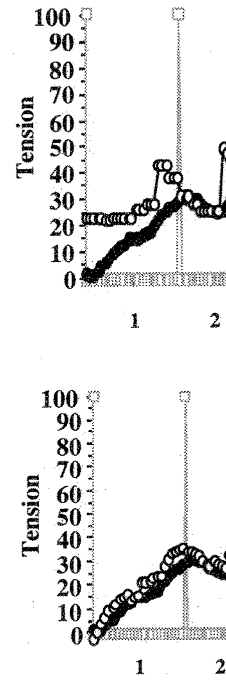

Fig. 9. The top grap
periments 1, 2, and 3
the model's variables
(1988) pitch-space t
the judgments. Com
information over tim
to the development o
beats), which produc
figure.

enths. These thre
The remaining va
chord is notated
ample, x = ii/E♭
major. *Pitch-spac*
between them al
the number of d
Pitch-space j dist
fifths for chords
number of tones
of the basic pitch
of counting weig
unshared tones:
for the third, an

value is the sum of the pitch space *i*, *j*, and *k* values for all events superordinate to an event in the prolongational tree. This gives a total of seven quantitative variables.

The fit of the tension judgments by the model with these seven variables was assessed by using multiple regression. The multiple regression was highly significant [$R(7,177) = .79$, $p < .0001$], indicating a good fit to the data. Subsequent regression analyses showed that the surface dissonance variables accounted for the data less well [$R(3, 181) = .28$, $p = .001$] than the four pitch-space variables [$R(4,180) = .78$, $p < .0001$]. Indeed, these four variables alone accounted for the data as well as all seven variables together. Of the pitch-space variables, the strongest was inherited value (see also Palmer, this issue), followed by pitch-space *k* distance.

The top of Figure 9 compares the tension judgments with the full seven-variable model. As can be seen, the responses tended to lag slightly behind the theoretical predictions. That is, a rise in predicted tension is often followed shortly by a rise in perceived tension and similarly for drops in tension. In addition, the model predicts more fine-grained detail than is found in the judgments. Together, these results suggest that listeners are integrating the musical information over time, producing the smoother and slightly delayed tension profiles. To test this, a model that included lags in units of 250 ms from 0 to 3250 ms (approximately two beats of the music) was tested. The fit of the model is shown at the bottom of Figure 9. This model with lags provided a considerably better fit to the data [$r(14,170) = .91$, $p < .0001$]. The smoothness of the tension judgments suggests that they would not be modeled well by either Narmour's (this issue, Figures 25–27) values of closure and nonclosure or the strength of Bharucha's (this issue) yearning vector, which appear to apply on a more local time scale.

Musical Ideas

Finally, Figure 10 shows the percentage of listeners identifying new musical ideas in each two-beat interval during the initial segment of the piece in Experiments 1, 2, and 3. Comparing first the results for Experiments 1 and 2, we see similar patterns. However, the responses tended to occur somewhat earlier in Experiment 2 than in Experiment 1. In part because of this, the correlation between Experiments 1 and 2 was negative [$r = -.23$ ($N = 57$, $df = 55$)], which approached significance, $p = .08$. However, shifting the Experiment 1 data earlier by three beats produced a good match between the experiments [$r = .75$ ($N = 57$, $df = 55$), $p < .0001$]. This temporal lag corresponds approximately to the time it takes for the new material in the melody and/or accompaniment to be repeated, which may be necessary in the first experiment to reinforce the impression that a new musical idea has been introduced. In addition to the temporal shift in responding in

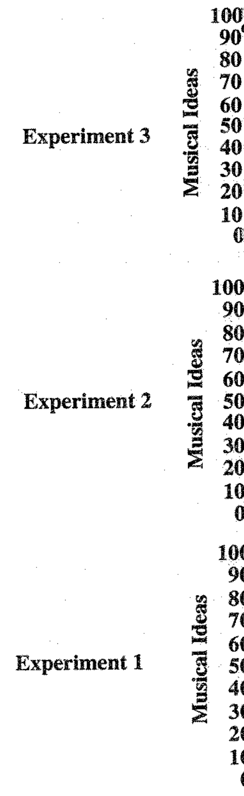

Fig. 10. Judgments of
sections produced res
idly in Experiments 2
Experiment 1. Compa
of new musical ideas
shows that new music

Experiment 2, mo
7, and 12 than be
Experiment 3 v
gave responses in
measure 2. These
[r = .71 (N = 32,
responses relative
negative correlati
and 3, which agai
by five beats proc
[r = .35 (N = 32,

Comparison with Figure 7 shows an increasingly close correspondence between the locations of judged endings and the introduction of new musical ideas. In the later experiments, judgments of section ends were always followed shortly by judgments of new ideas. These two variables had a correlation of .59 ($N = 57$, $df = 55$), $p < .0001$ in Experiment 2, and $r = .45$ ($N = 32$, $df = 30$), $p = .01$ in Experiment 3. This suggests that the smaller scale sections are defined largely by a figural strategy of segmentation. Comparison with Figure 8 shows the relationship between the locations of new musical ideas and tension was also stronger in these experiments. New musical ideas tended to be identified at points in the music where the tension level was either low or had just declined markedly. Indeed, the additional judgments of musical ideas that appeared in these later experiments in measures 2, 6, and 7 can be linked to drops in the tension values in these regions. Only those in measures 11 and 12 did not follow this pattern. Finally, new musical ideas tended to be introduced when tempo was at a neutral level. The correlations of these variables were .24 ($N = 57$, $df = 55$) and .05 ($N = 32$, $df = 30$) in Experiments 2 and 3, respectively, neither of which was statistically significant. In sum, on the smaller time scales, judgments of new musical ideas quite consistently followed section ending judgments and occurred at points of low tension and neutral tempo.

Experiment 4

Before turning to a discussion of the results, a fourth, subsidiary experiment will be presented. The first three experiments revealed a number of relationships between segmentation, tension, and musical ideas. In addition, some of these correlated with performed variations in tempo and dynamics. This raises the question as to the causal nature of the links between performance variations and perceptual responses. How would the responses change for a temporally regular or dynamically level performance? Listeners in the fourth experiment heard four different versions of the piece: as performed, constant dynamics (with performed tempo), constant tempo (with performed dynamics), and constant dynamics and tempo. In the interest of time, listeners made tension judgments only. This task was selected because it seemed intuitively to be the most susceptible to performance nuances. In contrast, segmentation and musical ideas would seem to be signaled by numerous cues contained in the notated pitches and durations independently of how the piece is performed.

METHOD

Subjects

Twenty-four members of the Cornell community participated in the experiment for which they received course credit. Listeners had from 1 to 18 years of instruction on musical

instruments, with a r
university level. Two
cording they owned,

Apparatus

Same as in Experi

Stimulus Material

All versions of th
measures 1–15. With
make a natural sounc
original performance
based on the perform
formed tempo), used
sponding to a moder
dynamics), used the
entire piece was the s
durations. The fourth
of the second version

Procedure

The display on the
ing the tension judgm
listener heard all four
each version was hear
questionnaire. The ex

The main focu
ers. Preliminary a
periment were lo
which, however,
correlations were
$r = .18$; constant
(performed dynar
Nor were there o
versions were pre
teners (and, conse
Figure 11 show
tain remarkably s
tween versions wi
performed versio
was .88 ($N = 897$
and third graphs
drop in tension at
lation between th

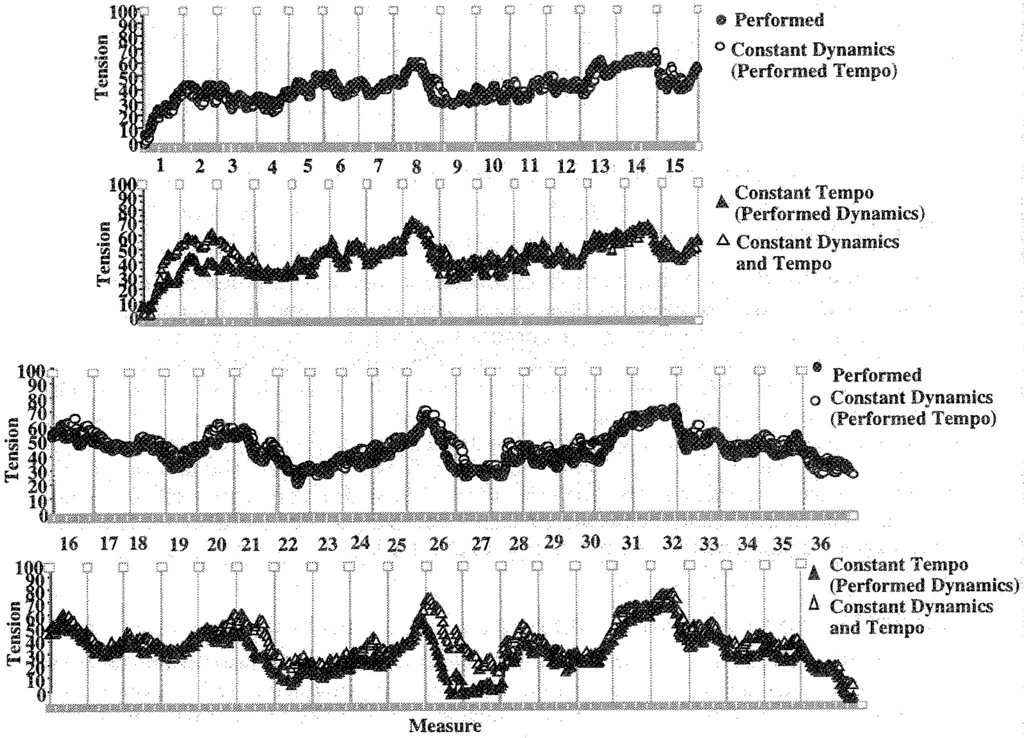

Fig. 11. Tension judgments from Experiment 4, which manipulated dynamics and tempo. Listeners heard four versions of the music: as performed, constant dynamics (with performed tempo), constant tempo (with performed dynamics), and constant dynamics and tempo. Only one repetition of measures 1–15 was played. The four tension curves showed remarkably similar patterns, with deviations apparent only at the ends of some of the major sections.

constant dynamics and tempo versions was $r = .83$ ($N = 897$, $df = 895$), $p < .0001$. These values are shown in the second and fourth graphs. These two differ primarily in measures 21 and 26, where the dynamics seem to enhance the large drop in tension. Comparison between the versions with different tempos is more difficult. However, visual inspection shows the graphs have very similar shapes, with nearly equal average values, ranges, and degrees of variation for all four versions. In general, it would seem that the manipulation of performed tempo and dynamics had remarkably little effect on the tension judgments.

Discussion

The discussion will focus on a number of issues of experimental methodology, and the primary empirical results will be reviewed in that context.

One of the metap
musical experienc
ously during the
phor was music a
of segmentation a
tion explored her
able responses wi
the musical mater
of music as it wou
elicited from liste
considerable relia
new musical idea
creased experienc
slowly the first tir

Another metho
of using a numbe
would the differe
addition, would
sively smaller sec
that held for the s
tent relationships

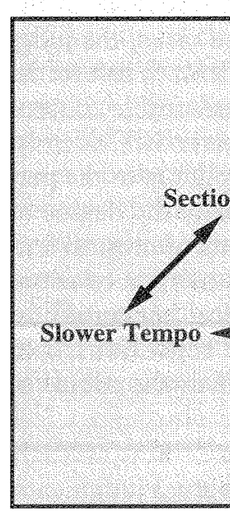

Fig. 12. Summary of c
section ends, tension
scale, judgments of se
smaller time scales, th
were preceded by jud
co-occurred with slov
co-occurred with neu

tion end had occurred were soon followed by judgments that a new musical idea had been introduced (and, conversely, judgments that a new musical idea had occurred were almost always preceded by judgments that a section end had occurred). Section-end responses corresponded to peaks in tension followed by rapid decreases in tension and slower tempos. In contrast, new musical ideas were introduced when tension was at a low level and the tempo was neutral. On the large scale, the relationships between these variables were not quite as one-to-one. In particular, listeners judged there to be more new musical ideas than major section ends. Consequently, judgments of section ends were always followed by judgments of new musical ideas but judgments of new musical ideas also occurred within large-scale segments. Employing the three tasks in combination showed these three aspects of music perception are generally quite tightly coordinated with one another.

Another methodological issue considered here was the influence of individual differences in musical experience. Does this cause listeners to respond differently from one another? Listeners in these experiments varied considerably in musical training, although none would qualify either as a professional musician or as a total novice. It is difficult to assess musical expertise and aptitude, but indicators such as years of formal instruction and extent of academic training showed considerable variability. Indeed, one of the listeners had, previously to the experiment, memorized and performed this particular sonata. Despite these differences, responses in the experiment were quite uniform. The nature of two of the tasks, the judgments of section ends and new musical ideas, made it difficult to test statistically the agreement between listeners. Nonetheless, considerable consensus was apparent, particularly about the locations of section ends. Evidently, these aspects of musical structure are expressed quite explicitly in the perceptual information. It was possible to examine statistically the degree to which listeners agreed with one another on their judgments of musical tension. Strong intersubject agreement was found, with no consistent relationship with musical training or other aspects of their musical backgrounds. Also, the tension judgments were highly reliable over repetitions, and changed little with increased experience with the piece over the course of the experiments.

Finally, the first three experiments raised the question of how strongly the performed dynamics and tempo influenced the perceptual judgments. These experiments showed that lower tension ratings tended to co-occur with lower dynamics and, as noted earlier, higher tension ratings tended to co-occur with the slower tempos. To examine this question, versions of the piece were created with constant dynamics and tension and were presented to listeners with the tension task. These manipulations produced remarkafter segment boundaries. This was particularly true for segmentation as judged on the smaller scale, where segmentation corresponded mostly closely to the introduction of new musical ideas.

The framework for discourse analysis proposed by Chafe (1994) brings out a number of other similarities. First, new musical ideas tended to be introduced at points of low tension and neutral tempo, which may correspond to his starting points or points of departure that are prepared by the larger context. Second, section ends identified by listeners at all levels in these experiments corresponded to slowing of tempo, perhaps analogous to the patterns of phrase final lengthening and pauses at the end of discourse units. Third, section ends, like the ends of prosodic units, tended to be marked by descending contour and decreased dynamics. Finally, the asymmetric patterns of tension within segments, noted above, may correspond in some way to how ideas are developed and completed within linguistic units.

Again, many unanswered questions remain. Do repetitions in music relate in some way to the units of semiactive information described by Chafe? Is there a quantifiable correspondence between the cognitive demands of musical and linguistic units? Their durations? Are comparable patterns found in other pieces, or other musical styles? Whatever the answers to these questions may be, the results of the experiments presented here suggest that this particular piece of music coordinates a number of different perceptual and conceptual structures in a way that invites comparison with linguistic discourse.

Comparisons between the experimental results and the theoretical analyses of the piece (Gjerdingen, this issue; Lerdahl, this issue; Narmour, this issue) also raise a number of questions. However, the numerous points of convergence suggest an increasing understanding of the musical structures that underlie music perception. One analysis by Gjerdingen that described how the piece divides into distinctive figures corresponded well to the new musical ideas identified by listeners. Although contemporary listeners are unlikely to have the associations to historical precedents described by Gjerdingen (this issue), they nonetheless appear able to identify the appropriate figural constituents. Narmour's (this issue) description of the formal design of the piece corresponded to listeners' segmentation judgments, and his qualitative analysis of sources of tension in the music corresponded to listeners' tension judgments. Finally, Lerdahl's (this issue) quantitative predictions of tonal tension provided a good model of the tension judgments in the opening section of the piece. The success of the model supported both local effects of harmonic tension and more global influences depending on an event's position in the proposed hierarchical tree.

Convergence of this sort with the perceptual data provides external validation for the experimental methods. In turn, the perceptual data help clarify

Clynes, M., & Nettheim, N. (1982). The living quality of music: Neurobiologic patterns of communicating feeling. In M. Clynes (Ed.), *Music, mind, and brain*. New York: Plenum.
Cook, N. (1987). *A guide to musical analysis*. New York: W. W. Norton.
Deliège, I, & El Ahmadi, A. (1990). Mechanisms of cue extraction in musical groupings: A study of perception on *Sequenza VI* for viola solo by Luciano Berio. *Psychology of Music, 18,* 18–44.
Fredrickson, W. (1995). A comparison of perceived musical tension and aesthetic response. *Psychology of Music, 23,* 81–87.
Gabrielsson, A. (1973). Adjective ratings and dimension analysis of auditory rhythm patterns. *Scandinavian Journal of Psychology, 14,* 244–260.
Hevner, K. (1936). Experimental studies of the elements of expression in music. *American Journal of Psychology, 48,* 246–268.
Hindemith, P. (1942). *The craft of musical composition*. Mainz: Schott.
Imberty, M. (1981). *Les écriture du temps*. Paris: Dunod.
Krumhansl, C. L. (1990). *Cognitive foundations of musical pitch*. New York: Oxford University Press.
Krumhansl, C. L. (1991). Music psychology: Tonal structures in perception and memory. *Annual Review of Psychology, 42,* 277–303.
Krumhansl, C. L. (1995). Music psychology and music theory: Problems and prospects. *Music Theory Spectrum, 17,* 53–80.
Kurth, E. (1947). *Musikpsychologie*. Bern: Verlag Krompholz. (Originally published 1931)
Lerdahl, F. (1988). Tonal pitch space. *Music Perception, 5,* 315–350.
Lerdahl, F., & Jackendoff, R. (1983). *A generative theory of tonal music*. Cambridge, MA: M.I.T. Press.
Lindblom, B., & Sundberg, J. (1970). Towards a generative theory of melody. *Swedish Journal of Musicology, 52,* 71–88.
Madson, C. K., & Fredrickson, W. E. (1993). The experience of musical tension: A replication of Nielsen's research using the continuous response digital interface. *Journal of Music Therapy, 30,* 46–63.
Meyer, L. B. (1956). *Emotion and meaning in music*. Chicago: University of Chicago Press.
Meyer, L. B. (1967). *Music, the arts, and ideas*. Chicago: University of Chicago Press.
Nattiez, J-J. (1990). *Music and discourse: Toward a semiology of music*. Princeton: Princeton University Press.
Nielsen, F. V. (1983). *Oplevelse af musikalsk spænding (The experience of musical tension)*. Copenhagen: Akademisk Forlag.
Palmer, C., & Krumhansl, C. L. (1987). Independent temporal and pitch structures in perception of musical phrases. *Journal of Experimental Psychology: Human Perception and Performance, 13,* 116–126.
Palmer, C., & Krumhansl, C. L. (1990). Mental representations for musical meter. *Journal of Experimental Psychology: Human Perception and Performance, 16,* 728–741.
Ratner, L. G. (1977). *Music: The listener's art* (3rd ed.). New York: McGraw-Hill.
Ratner, L. G. (1980). Classic music: *Expression, form, and style*. New York: Schirmer.
Ratner, L. G. (1991) Topical content in Mozart's keyboard sonatas. *Early Music, 19,* 615–619.
Repp, B. (1993). Music as motion: A synopsis of Alexander Truslit's (1938) *Gestaltung und Bewegung in der Musik*. *Psychology of Music, 21,* 48–72.
Rothfarb, L. A. (1991). *Ernst Kurth: Selected writings*. Cambridge: Cambridge University Press.
Sundberg, J. (1987). *The science of the singing voice*. Dekalb, IL: Northern Illinois University Press.
Sundberg, J., & Lindblom, B. (1976). Generative theories in language and music descriptions. *Cognition, 3,* 99–122.
Toth, E. (1977). *The shaping forces in music*. New York: Dover. (Original work published 1948)
Zuckerkandl, V. (1956). *Sound and symbol*. Princeton: Princeton University Press.

Music Perception
Spring 1996, Vol. 13, No. 3, 433–453

Anatomy of a Performance: Sources of Musical Expression

CAROLINE PALMER
Ohio State University

Research on music performance often assumes that a performer's intention to emphasize musical structure as specified in a score accounts for most musical expression. Relatively unstudied sources of expression in performance include notational variants of compositional scores, performer-specific aspects, and the cultural norms of a particular idiom, including both stylistic patterns that exist across musical works and expectations that arise from a particular musical context. A case study of an expert performance of a Mozart piano sonata is presented in which influences of historical interpretations of scores, performer-specific treatments of ornamentation and pedaling, and music-theoretic notions of melodic expectancy and tension-relaxation are revealed. Patterns of organization internal to the performance expression transcended the coarse-grained information given in scores, suggesting that performance is a better starting point than a musical score for testing theories of many musical behaviors.

DESPITE the fact that music performance provides a rich perspective on our musical experiences, our understanding of music performance has not caught up with empirical study of other types of musical experiences. Take for example the empirical research published in this journal (one of the leading journals in the study of music cognition) in the past 5 years: fewer than one fifth of the articles address performance, and twice as many address perception of music. Some of this disparity arises from the fact that only a minority of people perform music formally in our culture, whereas almost everyone listens to music. Another source of the disparity arises from the imperfect (or absent) methodologies for studying performance; a single performance typically results in a large quantity of complex measurements, and analysis techniques have been lacking. Recently, techniques for measuring and quantifying performance have improved with the advance of computer-aided musical instruments.

Requests for reprints may be sent to Caroline Palmer, Psychology Dept., 1885 Neil Ave., Ohio State University, Columbus OH 43210. (e-mail: cpalmer@magnus.acs.ohio-state.edu)