Sensitivity to combinations of musical parameters: Pitch with duration, and pitch pattern with durational pattern

WILLIAM FORDE THOMPSON
Atkinson College, York University, North York, Ontario, Canada

In four experiments, listeners' sensitivity to combinations of pitch and duration was investigated. Experiments 1–3 involved "textures" of notes, which were created by repeatedly sounding one of two notes (e.g., C_4 quarter note; D_4 eighth note), so that each note had an equal chance of occurring at any point within a texture. Experiment 1 showed that if a texture change was effected by introducing a pitch or duration that was not in the initial texture, the change was perceived by both attentive and distracted listeners. If a texture change was effected by combining the pitch of one note with the duration of the other note in the initial texture, and vice versa, it was perceived only if the listeners were attentive. Sensitivity to pitch/duration combinations was poorer when the pitch difference between component notes of textures was increased (Experiment 2), but it was better when the difference in duration between component notes was increased (Experiment 3). In Experiment 4, listeners' sensitivity to combinations of pitch pattern and durational pattern in brief sequences was examined. Listeners were sensitive to the manner in which parameter patterns were combined when they were attentive, but not when they were distracted. The results are discussed in view of feature-integration theory and its application to music cognition.

In this investigation, I examined listeners' sensitivity to the manner in which pitch and duration are combined.[1] Two theoretical issues are relevant to the present research. The first issue concerns the study of music cognition, which may be informed by examining listeners' sensitivity to combinations of musical parameters. The second issue concerns "feature-integration theory" and its application to the qualities of pitch and duration. These issues will be reviewed in turn.

Combinations of Musical Parameters

In psychological studies of music, musical parameters are often examined one at a time. For example, music research may focus on sensitivity to pitch-related information (A. J. Cohen, 1991; Cuddy & Thompson, 1992; Krumhansl, 1990; Thompson & Mor, 1992) or temporal structure (e.g., Desain, 1992; Gabrielsson, 1973a, 1973b; Longuet-Higgins & Lee, 1982). However, many researchers have become interested in the effects of combining musical parameters.

One way to address this issue is to examine whether the effect of one parameter is dependent on the properties of another parameter. For example, Jones, Boltz, and

Kidd (1982) demonstrated that listeners are better able to recognize the pitch of a tone when it is rhythmically accented than when it is not rhythmically accented. Conversely, when pitch structure and temporal structure do not imply compatible groupings, memory is impaired (Boltz & Jones, 1986; Deutsch, 1980; Jones, 1987). These studies show that pitch and temporal structure have joint influences on judgments.

A second way to address the issue of combining parameters is to assess whether the influences of two or more parameters are statistically independent. Studies have suggested that pitch pattern and temporal pattern make independent contributions to judgments of melodic similarity (Monohan & Carterette, 1985) and phrase completion (Palmer & Krumhansl, 1987a, 1987b). Findings of statistical independence are compatible with the possibility that parameters are processed independently. However, independence of processing is not the only interpretation of statistical independence. Statistical independence merely indicates that combining parameters has not resulted in emergent qualities that influence judgments beyond what can be explained by the individual parameters.

A third way to address the issue of combining parameters is to assess the degree to which listeners are sensitive to combinations of parameters. Krumhansl (1991) examined listeners' sensitivity to combinations of parameters in a short atonal piano composition by Messiaen. The piece was composed so that each pitch was associated with a specific duration. Listeners were presented the first half of the piece, followed by a second segment

of music. Their task was to judge whether the first and second segments of music belonged to the same piece. The judgments were not influenced by whether pitches and durations were combined similarly in the two segments, suggesting that the listeners were not sensitive to this aspect of musical structure. As Krumhansl has pointed out, the processing implications of the finding are ambiguous. One possibility is that the listeners were unable to encode the pitch/duration combinations (e.g., because they are processed independently). However, it is also possible that they simply did not consider the possibility that such combinations were relevant to their judgments.

In the present investigation, a similar approach was used to examine the issue of combining parameters. However, listeners' ability to perceive and remember combinations of pitch and duration was more directly assessed. Specifically, the experiments were assessments of listeners' ability to perceive changes in the way pitch and duration are combined.

Feature-Integration Theory

The design for the present research was based on experiments used to test feature-integration theory in vision (Treisman & Gelade, 1980, Experiments 5–7), and feature-integration theory provided the theoretical framework for the research. The following is an overview of feature-integration theory, including examples of research that relate this theory to musical phenomena.

One of the basic assumptions of feature-integration theory is the idea of automatic feature detection in the initial stages of processing. In vision, various lines of evidence suggest that visual qualities of objects, such as color, size, orientation, and direction of movement, are analyzed automatically and independently (Livingstone & Hubel, 1987; Treisman, 1986). In audition, independent analysis of qualities may also occur. Pitch, duration, timbre, intensity, and location are strong candidates for being distinct and independent aspects of the acoustic sensory mechanism. However, research on the analysis of auditory features has been given less attention than research on the processing of visual features.

It is believed that a process of feature integration occurs after the feature-detection stage. Feature integration depends on information-processing conditions more heavily than does feature analysis. Accurate integration of features is assumed to occur under optimal circumstances. However, inaccurate integration of features may occur under nonoptimal conditions—for example, when a person is distracted, when stimulus dimensions are not easily discriminable, or when perceptual organization conflicts with proper integration (A. Cohen & Ivry, 1989; Ivry & Prinzmetal, 1991; Treisman & Gelade, 1980). The perceptual effects of inaccurate feature integration are referred to as *illusory conjunctions* (Treisman & Schmidt, 1982).

Deutsch (1975) reported effects for musical stimuli that are analogous to the illusory conjunctions that have been reported in visual studies (see also Butler, 1979; Hall & Pastore, 1993). Under certain conditions, listeners do not accurately perceive the manner in which location and pitch information are combined. Two major scales, one ascending and the other descending, were presented simultaneously to listeners. Successive tones in each of the two scales were presented alternately to the left and right ears, so that when a tone from the ascending scale was presented to the right ear, a tone from the descending scale was presented to the left ear. The manner in which locations and pitches were combined was not accurately perceived by the listeners. Rather, locations were combined with pitches in a way that preserved pitch proximity. This effect, called the *scale illusion*, indicates that listeners are not always sensitive to the manner in which pitch and location are combined.

Although there have been no reports of illusory conjunctions for pitch and duration, Peretz and her associates have found evidence that pitch and temporal information are neurally separated in the central nervous system (Peretz & Kolinsky, 1993; Peretz & Morais, 1989). These reports document conditions of brain damage in which patients suffer from a deficit in temporal processing with no deficit in pitch processing, or vice versa. One patient, who sustained a lesion in the left temporal lobe, was unable to discriminate between sequences differing only in temporal structure, but he had no difficulty in discriminating between sequences differing in pitch structure. A patient with damage to the right temporal lobe showed the reverse effect: discrimination on the basis of pitch structure was impaired, but discrimination on the basis of temporal structure remained at normal performance levels. According to Peretz and Morais, these deficits point to specific neural modules for pitch and temporal information and reveal a neural dissociation between pitch and temporal processing.

If temporal information is neurally dissociated from pitch information, then it is possible that normal listeners may not always be sensitive to the manner in which these parameters are combined. This idea was examined by Thompson and Sinclair (1993). Pairs of melodic sequences were played either consecutively or simultaneously, and listeners were asked to make judgments requiring sensitivity to either individual parameters or combinations of parameters. The results showed that the listeners were not highly sensitive to the manner in which pitch patterns were combined with durational patterns. This suggests that neural dissociation between pitch pattern and durational pattern is functionally manifested in a partial insensitivity to the manner in which these parameters are combined.

In Experiments 1–3, I examined listeners' ability to perceive changes in how pitch and duration are combined in individual notes. The experiments involved an auditory texture-segregation task, derived from Treisman and Gelade's (1980) visual texture-segregation tasks. In Experiment 4, a modified task was used to examine listeners' sensitivity to the manner in which pitch

pattern is combined with durational pattern in short melodies. If perceptual processing involves encoding combinations of pitch and duration—that is, perceptual integration—then listeners should readily perceive changes in how pitch and duration are combined. If pitch and duration are processed independently, or if perceptual integration is incomplete, they should have difficulty perceiving these changes.

EXPERIMENT 1

The design of Experiment 1 was based on Treisman and Gelade's (1980, Experiments 5–7) texture-discrimination task, which they used to evaluate feature-integration theory. In the present experiment, an auditory analogue of the task was used to assess listeners' sensitivity to combinations of pitch and duration. A *note texture* was operationally defined as a monophonic note sequence involving two notes that alternated or repeated in a random fashion. Presentations consisted of either a single note texture, or a note texture that changed to another note texture. The listeners indicated whether or not they perceived a change in texture.

There were four types of texture presentations—three involved a texture change and one did not. The conditions involving a texture change are referred to as *simple-change*, *disjunct*, and *switch*. In the simple-change condition, the texture change was effected by introducing a feature that was not in the initial texture—either a new pitch or a new duration. In the disjunct condition, the texture change was effected by introducing both a new pitch and a new duration. In the switch condition, the initial and final textures contained the same pitch and duration values, but the manner in which these qualities were combined was different. For example, if the initial texture consisted of C_4 quarter notes and D_4 eighth notes, then the final texture consisted of C_4 eighth notes and D_4 quarter notes.

To perceive a texture change for simple-change or disjunct conditions, it was necessary for the listeners to perceive only the individual parameters involved; that is, perceiving a texture change for these conditions did not require sensitivity to the manner in which pitch and duration were combined. To perceive a texture change for switch conditions, however, it was necessary for the listeners to be sensitive to the manner in which pitch and duration were combined.

A key component of feature-integration theory is that focused attention plays a critical role in facilitating accurate integration of visual features (Briand & Klein, 1989; Treisman, 1986, 1991; Treisman & Gelade, 1980; Treisman & Sato, 1990). Therefore, the importance of attention was also examined. The listeners performed the task under conditions of focused attention and distraction.

Three predictions were assessed. Prediction 1 was that attentive listeners would be able to perceive the texture change in the switch condition; that is, they would be sensitive to combinations of parameters. Prediction 2 was that distracted listeners would be unable to perceive texture changes in switch conditions. Predictions 1 and 2 were based on Treisman and Gelade's (1980) view that focused attention is necessary for accurate perceptual integration of features in unfamiliar contexts. Thus, attentive listeners, but not distracted listeners, should be sensitive to the manner in which pitch and duration are combined. Prediction 3 was that texture changes would be easier to perceive in simple and disjunct conditions than in switch conditions, for conditions of both attention and distraction. This prediction was based on the idea that perceiving texture changes for simple and disjunct conditions relies on only the analysis of individual parameters, whereas perceiving texture changes for switch conditions relies on the integration of parameters. Effects arising from the analysis of individual parameters are predicted to be highly salient, because this process is thought to occur automatically and preattentively (Treisman & Gelade, 1980). Effects that rely on an integration process, however, are predicted to be less conspicuous, because integration is thought to depend more heavily on information-processing conditions (e.g., degree of attention) than does the analysis of individual parameters (Ivry & Prinzmetal, 1991; Treisman, 1986, 1991).

Method

Listeners. Fifteen undergraduate students at York University participated. The listeners had completed an average of approximately 5 years of formal training in music (range of 0 to more than 10 years). They were paid $7.00 for their participation, which included participation in Experiments 2 and 3. All the listeners reported having normal hearing.

Stimuli and conditions. The stimuli were generated with a Roland U-20 multitimbral keyboard with equal-temperament tuning, which was controlled by a Macintosh SE-30 computer. Each presentation consisted of 40 notes, with no pauses between them. The timbre of all of the notes was set to the Piano 1 sound of the Roland U-20 (for an acoustic description, see Fletcher & Rossing, 1991, chap. 12.8). The presentations involved either one single note texture, or a note texture that changed to another note texture. A single note texture was defined by two notes, each with a specified pitch and duration. For each note texture, the two defining notes were played repeatedly and monophonically, in an order that was determined randomly with replacement.

There were four texture conditions: no-change condition, simple-change condition, disjunct condition, and switch condition. There were four no-change textures, two simple-change textures, two disjunct conditions, and two switch textures. The component notes that were used to create the initial and final texture for the no-change condition, simple-change condition, disjunct condition, and switch condition are shown in Figure 1.

In the no-change condition, each of the 40 notes of the texture was one of the two notes that defined a single note texture. In the other conditions, a "texture change" occurred between Notes 16 and 24; the exact point of the texture change was determined randomly and independently for each trial involving a texture change. There was no pause between the initial and final textures for these conditions. In the first example of the simple-change condition, the texture change was effected by introducing a novel pitch value in the final texture—that is, by assigning the notes in the final texture a pitch value that had not occurred in the initial texture. In the
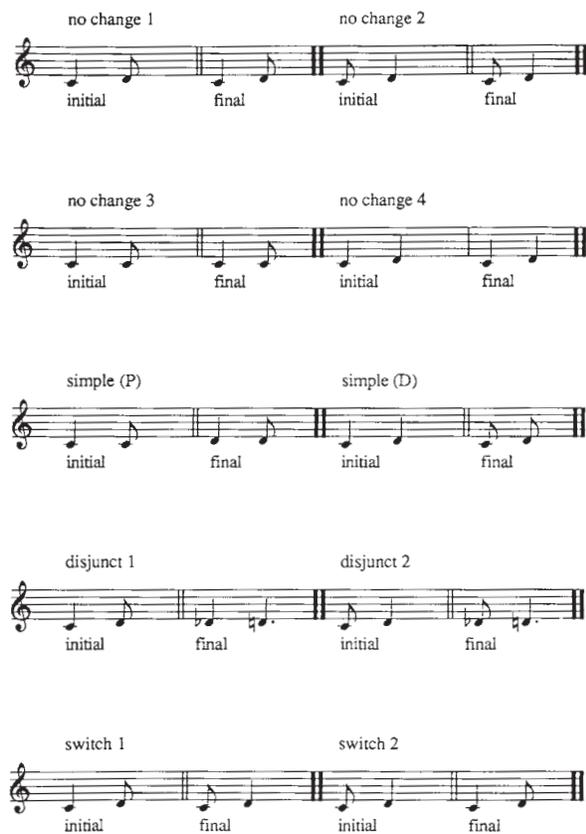
Figure 1. Component notes used in Experiment 1 to create the initial and final texture for each of the 10 presentations. (P, pitch; D, duration.)

second example of the simple-change condition, the texture change was effected by introducing a novel duration value in the final texture. In the disjunct condition, the texture change was effected by introducing a novel pitch value and a novel duration value in the final texture. In both examples of the switch condition, the texture change was effected by combining the pitch value of the first note in the initial texture with the duration value of the second note in the initial texture, and vice versa.

The two notes that were used to define note textures for the no-change condition were derived from the two initial textures for the simple-change condition and from the two initial textures for the switch condition (which are the same as the two initial textures for the disjunct condition). In the no-change condition, however, the initial texture continued until all 40 notes were sounded.

As is shown in Figure 1, the pitches of component notes were selected from C_4 and D_4, with $D\flat_4$ also used for the disjunct condition. Component notes for the no-change, simple-change, and switch conditions had duration values selected from 100 and 200 msec. The total duration of each presentation varied slightly, depending on the manner in which the two notes of a given texture were randomly sounded (except for the fourth no-change presentation, in which all 40 notes were 200 msec in length). The mean duration of these presentations was approximately 6.25 sec. Component notes for disjunct conditions had duration values selected from 100 and 200 msec in the initial texture, and 100 and 300 msec in the final texture. The mean duration of these presentations was approximately 7.0 sec.

Each note texture was presented twice to the listeners—once while they were attentive and once while they performed a dis-

traction task (described below). There were 20 trials in total, and they were presented to the listeners in a random order that was determined independently for each listener.

Distraction task. For half of the presentations, the listeners performed a distraction task while the note texture was being presented. The distraction task ensured that the listeners could not focus their attention on the texture presentation. In the distraction task, words appeared on the computer screen (e.g., *moon*, *eat*, *orange*, *run*), and the listener indicated, by clicking the mouse on one of two boxes, whether the word was most commonly used as a verb or a noun. If a response was not made within 2 sec, it was counted as incorrect, and the next word appeared. The distraction task began 1.0 sec before the test sequence was presented, and it stopped when the presentation ended. The number of correct responses on the distraction task was recorded, and it was displayed briefly to the listeners.

Procedure. The operational definition of a note texture was described to each listener. A change in texture was defined as any change in one or both of the two component notes that defined a texture. Following each presentation, a response prompt appeared on the computer screen, and the listeners made their responses directly on the computer. They indicated whether there was or was not a change in texture, and they rated the confidence of their decision on a scale from 1 (*confident*) to 3 (*not confident*). Practice trials, selected randomly from the set of experimental trials, were provided to acquaint the listeners with the task. They were allowed to quit the practice trials and begin the experimental trials at any time.

Results

The responses were first coded on a 6-point scale, in which 1 indicated that the listener was very confident that there was no change in texture, and 6 indicated that the listener was very confident that there was a change in texture. Table 1 displays the mean score for each condition in Experiments 1–3.

For the main analysis, scores for each no-change condition were subtracted from scores for the corresponding change condition (i.e., the change condition with the same initial texture). These "difference" scores (change minus no-change) account for response bias and provide a measure of the listeners' sensitivity to the different types of texture change. Difference scores for switch conditions, simple conditions, and disjunct conditions were subjected to an analysis of variance (ANOVA) with repeated measures on three factors: attention (2 levels), type of change (3 types—switch, simple, and disjunct), and example (2 examples of each type). In consideration of the assumption of compound symmetry, significant effects were confirmed with Greenhouse-Geisser and Huynh-Feldt corrections.

The ANOVA revealed a main effect of attention $[F(1,14) = 17.04, p < .005]$ and a main effect of type of change $[F(2,14) = 27.09, p < .0001]$. There was no significant interaction between attention and type of change $[F(1,14) = .07, p = .93]$. Figure 2 illustrates mean difference scores for the switch conditions and the simple/disjunct conditions, for the attention and distraction conditions. Simple and disjunct conditions are collapsed in the figure because there was no significant difference between these conditions, either for conditions of attention $[F(1,14) = .03, p = .87]$ or for distraction $[F(1,14) =$

Table 1
Mean Scores for Each Condition in Experiments 1–3, With Values for Switch Conditions, Simple Changes in Pitch (P), Simple Changes in Duration (D), and Disjunct Conditions

Condition	Experiment 1		Experiment 2		Experiment 3	
	Attention	Distraction	Attention	Distraction	Attention	Distraction
Switch						
Change	3.83	2.93	2.90	2.90	4.83	3.90
No change	2.10	2.73	2.17	2.50	2.27	2.57
Simple (P)						
Change	5.27	4.60	4.80	4.67	5.13	4.33
No change	1.93	2.47	1.47	1.73	1.47	1.93
Simple (D)						
Change	5.60	4.80	5.53	5.07	5.87	3.47
No change	1.33	2.07	2.47	2.87	2.00	2.47
Disjunct						
Change	5.83	5.17	6.00	5.40	5.97	5.30
No change	2.10	2.73	2.17	2.50	2.27	2.57

$0.00, p = 1.00$]. In addition, there was no significant difference between simple changes in pitch and simple changes in duration, either for conditions of attention [$F(1,14) = 3.6, p = .08$] or for distraction [$F(1,14) = 1.82, p = .20$].

The three basic predictions were assessed with planned comparisons. To test Predictions 1 and 2, the change and no-change scores for the switch condition were uncoupled and compared (see Table 1 for means). First, when the listeners were attentive, the mean score for the switch (change) condition was significantly higher than the mean score for the corresponding no-change condition [$F(1,14) = 24.33, p < .0005$]. This effect indicates that the attentive listeners were sensitive to the manner in which pitch and duration were combined. Second, when the listeners were distracted, the mean score for the switch condition was not significantly higher than the mean score for the corresponding no-change condition [$F(1,14) = 0.19, p = .67$]. Together, the latter two findings suggest that attention enhances sensitivity to combinations of pitch and duration. A comparison of the difference scores for the attention and distraction conditions of the switch conditions confirmed that this influence of attention was reliable [$F(1,14) = 8.34, p < .02$].

Prediction 3 was also supported by a planned comparison of difference scores. The mean difference score for the simple and disjunct conditions was significantly higher than the mean difference score for the switch condition, for attentive listeners [$F(1,14) = 18.53, p < .001$] and distracted listeners [$F(1,14) = 31.68, p < .0005$]. This finding is compatible with the prediction that effects involving the analysis of individual parameters are more conspicuous than those relying on the integration of parameters.

EXPERIMENT 2

The results of Experiment 1 support the three predictions. When listeners were attentive, but not when they were distracted, they were sensitive to the manner in which pitch and duration were combined. However, sensitivity to combinations of parameters was poorer than sensitivity to individual parameters. These effects were further examined in Experiment 2 for different stimulus conditions. In this case, texture changes were created with notes separated by a wide pitch interval. The pitches used in Experiment 2 were C_4, G_4, and E_5.

In general, the predictions that were made in Experiment 1 were also made in Experiment 2. However, there was the possibility that increasing the discriminability of component pitches might influence sensitivity to combinations of duration and pitch. Ivry and Prinzmetal (1991), for example, found fewer illusory conjunctions of color and shape between items with dissimilar colors or shapes than between items with similar colors and shapes (see also Garner & Felfoldy, 1970). In other words, integration of features was more accurate when items in the context were made less similar. Thus, increasing the pitch difference between notes may decrease their perceived similarity, facilitating the integration of pitch and duration and leading to improved
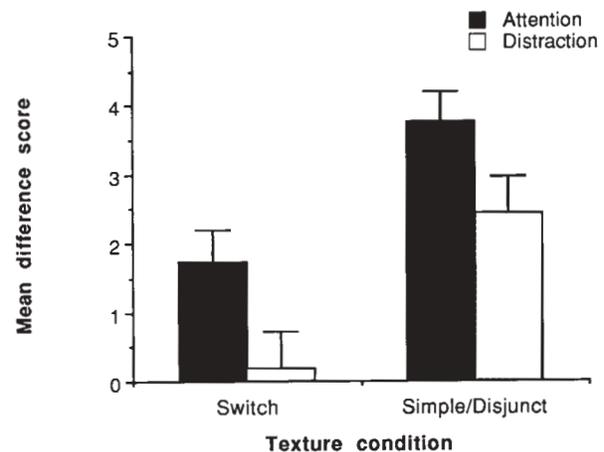
Figure 2. Mean difference scores for the switch condition and the simple-change/disjunct condition, for the attention and distraction conditions of Experiment 1.

sensitivity to the manner in which these parameters are combined.

Method

The participants, conditions, and procedure were identical to those described in Experiment 1. Component notes again had duration values of 100, 200, and 300 msec, but for this experiment they had pitch values selected from C_4, G_4, and E_5. Presentation conditions were identical to those used in Experiment 1, except that instances of D_4 were replaced by E_5, and instances of Db_4 were replaced by G_4 (see Figure 1).

Results

The responses were again coded on a 6-point scale; the means for each condition are displayed in Table 1. From these scores, difference scores were obtained by subtracting the scores for the no-change conditions from the scores for the corresponding change conditions. Difference scores for switch conditions, simple conditions, and disjunct conditions were subjected to an ANOVA with repeated measures on three factors: attention (2 levels), type of change (3 types), and example (2 examples). Significant effects were confirmed with Greenhouse-Geisser and Huynh-Feldt corrections.

There was no significant effect of attention [$F(1,14) = 3.52$, $p = .08$], but there was a main effect of type of change [$F(2,14) = 31.71$, $p < .0001$]. There was no significant interaction between attention and type of change [$F(1,14) = .71$, $p = .50$]. Figure 3 illustrates the mean difference scores for the switch conditions and the simple/disjunct conditions, for the attention and distraction conditions. The simple and disjunct conditions were again collapsed, because there was no significant difference between these conditions for conditions of attention [$F(1,14) = 1.04$, $p = .33$] or distraction [$F(1,14) = 1.43$, $p = .25$]. In addition, there was no significant difference between simple changes in pitch and simple changes in duration, for conditions of attention [$F(1,14) = .16$, $p = .69$] or distraction [$F(1,14) = 2.22$, $p = .16$].

The three basic predictions were assessed with planned comparisons. First, when the listeners were attentive, the mean score for the switch (change) condition was significantly higher than the mean score for the corresponding no-change condition [$F(1,14) = 5.81$, $p < .05$; see Table 1]. This effect suggests that the attentive listeners were sensitive to the manner in which pitch and duration were combined. Second, when the listeners were distracted, the mean score for the switch condition was not significantly higher than the mean score for the corresponding no-change condition [$F(1,14) = 1.86$, $p = .19$]. The latter two findings imply that attention enhances sensitivity to combinations of pitch and duration. However, a comparison of the difference scores for the attention and distraction conditions of the switch conditions failed to confirm that this influence of attention was reliable [$F(1,14) = .87$, $p = .37$].

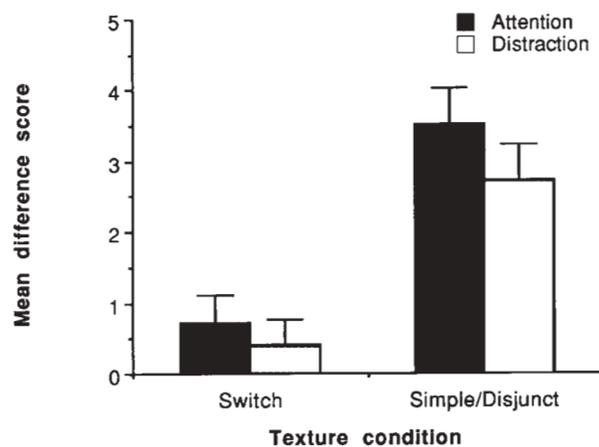Prediction 3 was again supported by a planned comparison of difference scores. The mean difference score

Figure 3. Mean difference scores for the switch condition and the simple-change/disjunct condition, for the attention and distraction conditions of Experiment 2.

for simple and disjunct conditions was significantly higher than the mean difference score for the switch condition when listeners were attentive [$F(1,14) = 64.79$, $p < .0001$] and when they were distracted [$F(1,14) = 37.38$, $p < .0001$]. This finding suggests that effects resulting from the analysis of individual parameters are more salient than those requiring sensitivity to combinations of parameters.

EXPERIMENT 3

Experiment 2 confirmed that when textures are composed of notes separated by a wide pitch interval, attentive listeners are still sensitive to the manner in which pitch and duration are combined. However, a comparison of Experiments 1 and 2 indicates that sensitivity to combinations of pitch and duration was poorer when notes were separated by a large pitch interval than when notes were separated by a small pitch interval. It is possible that the conspicuously large pitch differences diverted the listeners' attention away from duration information, resulting in a reduced sensitivity to the manner in which duration and pitch were combined.

Experiment 3 was a replication of Experiment 1, but the difference between durations used in texture presentations was increased. Here, textures were created with notes that had durations of 100 and 400 msec for each condition. The pitches used were C_4 and D_4. If diverting attention away from duration information can disrupt sensitivity to combinations of pitch and duration, then exaggerating the difference between durations in textures should have the reverse effect of facilitating sensitivity to combinations of pitch and duration.

Method

The participants, conditions, and procedure were identical to those described in Experiment 1.[2] Component notes had duration values of 100, 250, and 400 msec and pitches of C_4, Db_4, and D_4. Presentation conditions were identical to those used in Experi-

ment 1, except that instances of 200-msec notes were replaced by 400-msec notes, and instances of 300-msec notes were replaced by 250-msec notes (see Figure 1).

Results

The responses were again coded on a 6-point scale. The means for each condition are shown in Table 1. Difference scores were obtained by subtracting the scores for the no-change conditions from the scores for the corresponding change conditions. Difference scores for switch, simple, and disjunct conditions were subjected to an ANOVA with repeated measures on three factors: attention, type of change, and example. Significant effects were confirmed with Greenhouse-Geisser and Huynh-Feldt corrections.

There was a main effect of attention [$F(1,14) =$ 10.12, $p < .01$] and a main effect of type of change [$F(2,14) = 9.70$, $p < .001$]. There was no significant interaction between attention and type of change [$F(1,14) = .18$, $p = .84$]. Figure 4 shows the mean difference scores for the switch conditions and the simple/disjunct conditions, for the attention and distraction conditions. The scores for the simple and disjunct conditions are collapsed because they were not significantly different for conditions of attention [$F(1,14) = .03$, $p = .87$] or distraction [$F(1,14) = .05$, $p = .83$]. Scores for simple changes in pitch and duration were not significantly different for conditions of attention [$F(1,14) = .28$, $p = .61$] or distraction [$F(1,14) = .44$, $p = .52$].

The three predictions were assessed with planned comparisons. The mean score for the switch (change) condition was significantly higher than the mean score for the corresponding no-change condition, when the listeners were attentive [$F(1,14) = 22.34$, $p < .0005$] and when they were distracted [$F(1,14) = 15.34$, $p < .002$; see Table 1]. These effects indicate that the listeners were sensitive to the manner in which pitch and duration were combined, regardless of whether they were atten-
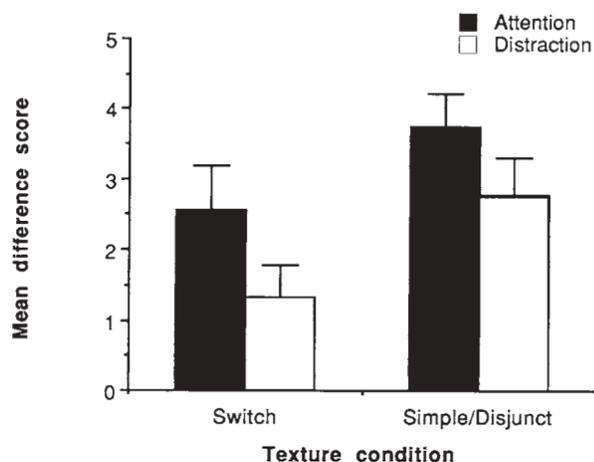

Figure 4. Mean difference scores for the switch condition and the simple-change/disjunct condition, for the attention and distraction conditions of Experiment 3.

tive. In other words, sensitivity to combinations of pitch and duration did not require focused attention when durations were distinctive. Nonetheless, a comparison of the difference scores for the attention and distraction conditions of the switch conditions indicated that attention significantly enhanced sensitivity to combinations of parameters [$F(1,14) = 5.42$, $p < .05$].

Support was again found for Prediction 3. The mean difference score for simple and disjunct conditions was significantly higher than the mean difference score for the switch condition when listeners were attentive [$F(1,14) = 8.51$, $p < .02$] and when they were distracted [$F(1,14) = 13.68$, $p < .005$]. This finding again suggests that effects resulting from the analysis of individual parameters are more salient than those requiring sensitivity to combinations of parameters.

EXPERIMENT 4

Experiments 1–3 showed that, for varying stimulus conditions, attentive listeners are sensitive to the manner in which pitch and duration are combined in individual notes. In general, attention facilitated sensitivity to combinations of pitch and duration. Experiment 3, however, shows that even distracted listeners are sensitive to combinations of pitch and duration when durations are highly distinct. Finally, a comparison of judgments for the simple-change and switch conditions suggests that effects requiring sensitivity to individual parameters are more pronounced than those requiring sensitivity to combinations of parameters.

Extending the approach taken in Experiments 1–3, sensitivity to combinations of pitch and durational patterns in brief melodic sequences was examined in Experiment 4 (see also Thompson & Sinclair, 1993). Another reason for conducting Experiment 4 was to assess the possibility that the effects reported in Experiments 1–3 were limited to artificial or unfamiliar stimuli, such as "note textures." Experiment 4 involved simple scalar pitch movements and simple rhythms, which are typical in Western music and familiar to listeners.

A modified procedure was adopted. Listeners were presented with two short note sequences, followed by a pause, and then two more note sequences. The first two sequences were called the *test sequences*, and the second two sequences were called the *comparison sequences*. The test and comparison sequences were either identical, or differed in one of three ways. The listeners were asked to judge whether or not there was a difference between test and comparison sequences.

The differences between the test and comparison sequences were comparable to the texture-change conditions used in Experiments 1–3. In the simple-change and disjunct conditions, the comparison sequences involved a pitch pattern and/or durational pattern that was not in the test sequences. In the switch condition, the test and comparison sequences involved the same pitch patterns and durational patterns, but they differed in how these

pitch patterns and durational patterns were combined. That is, Comparison Sequence 1 was created by combining the pitch pattern of Test Sequence 1 with the durational pattern of Test Sequence 2. And Comparison Sequence 2 was created by combining the pitch pattern of Test Sequence 2 with the durational pattern of Test Sequence 1. Again, the role of attention was examined, and the three predictions were assessed.

Prediction 1 was that attentive listeners would be sensitive to the difference between test and comparison sequences in the switch condition. That is, attentive listeners would be sensitive to combinations of pitch pattern and durational pattern. Prediction 2 was that distracted listeners would not be sensitive to the difference between test and comparison sequences in the switch condition. That is, they would be insensitive to the manner in which pitch pattern and durational pattern were combined. Predictions 1 and 2 were based on Treisman and Gelade's (1980) view that focused attention is necessary for accurate perceptual integration. Prediction 3 was that, for both attention and distraction conditions, the listeners would be more sensitive to the difference between test and comparison sequences for simple-change and disjunct conditions than for switch conditions. This prediction was based on the idea that effects relying only on sensitivity to individual parameters should be highly salient, because feature analysis occurs automatically and preattentively. Effects relying on sensitivity to combinations of parameters should be less salient, because the integration of parameters depends more heavily on information-processing conditions (e.g., degree of attention) than does feature analysis.

Method

Listeners. Fifteen adult listeners participated in the experiment. They were at least moderately trained in music, with a minimum of 5 years of formal training on a musical instrument. All the listeners reported having normal hearing.

Stimuli and conditions. Sequences were generated with a Roland U-20 multitimbral keyboard with equal-temperament tuning, which was controlled by a Macintosh SE-30 computer. Each presentation consisted of four consecutive five-note sequences. The timbre of all the notes was set to the Piano 1 sound of the Roland U-20. The first two sequences were the test sequences, and the last two sequences were the comparison sequences. The order in which the two test sequences occurred and the order in which the two comparison sequences occurred were randomly determined for each listener. Each sequence was randomly transposed so that its starting note was between C_4 and G_4. Randomizations of transpositions and sequence order were determined independently for each listener.

The tempo of the sequences was 200 quarter-tone beats/min. There was a pause equal to 1 beat between the first and second sequences and between the third and fourth sequences. There was a pause of 1.0 sec between test and comparison sequences.

The listeners' task was to compare the test sequences with the comparison sequences, without regard to order of occurrence or pitch transposition. The two test sequences were always different from each other. In four presentations, the two comparison sequences were the same as the two test sequences, disregarding order of occurrence and pitch transposition (no-change conditions). In four other presentations, the comparison sequences were different from the test sequences, as described below (change conditions). The subjects listened to each of the eight presentations twice—once while performing a distraction task and once while undistracted. Presentation order for the 16 trials was randomized independently for each listener.

Figure 5 displays the two test sequences and the two comparison sequences for each of the four presentations in which test and comparison sequences differed. These four presentations represent the following conditions: simple (PP, or pitch pattern) change condition, simple (DP, or durational pattern) change condition, disjunct condition, and switch condition. In the simple (PP) change condition, the test sequences shared the same pitch pattern, and the comparison sequences had a novel pitch pattern (i.e., a pitch pattern that was different from that of the test sequences). In the simple (DP) change condition, both test sequences shared the same durational pattern, and the comparison sequences had a novel durational pattern. In the disjunct condition, the comparison sequences were created by combining the durational pattern of one sequence with a novel pitch pattern, and by combining the pitch

Figure 5. Sequences used in Experiment 4. The figure illustrates the two test sequences and the two comparison sequences for each of the four presentations in which test and comparison sequences differed. (PP, pitch pattern; DP, durational pattern.)

pattern of the other sequence with a novel durational pattern. In the switch condition, the two comparison sequences were created by combining the durational pattern of one test sequence with the pitch pattern of the other test sequence, and vice versa. That is, the comparison sequences of the switch conditions contained no novel pitch patterns or durational patterns, but they differed from the test sequences with respect to how the pitch and durational patterns were combined.

For the four no-change conditions, the two test sequences were the same as the two comparison sequences. These sequences were also the same as the following sequences used in the change conditions: (1) the test sequences for the simple (PP) change condition, (2) the comparison sequences for the simple (DP) change condition, (3) the comparison sequences for the disjunct condition, and (4) the test sequences for the switch condition.

Distraction task. The distraction task was identical to that used in Experiment 1.

Procedure. Following each presentation, the listeners indicated whether or not the first pair of sequences was the same as the second pair of sequences by responding "yes" or "no" to a prompt that appeared on the computer screen. A difference between the two pairs was defined as any difference between the two test sequences and the two comparison sequences other than order of presentation or transposition in pitch. Practice trials, selected randomly from the set of experimental trials, were provided to acquaint the listeners with the task. The listeners were allowed to quit the practice trials and begin the experimental trials at any time.

Results

Responses were recorded as "1" (test and comparison were different) and "0" (test and comparison were the same). Difference scores were then obtained by subtracting no-change values from values for the corresponding condition in which test and comparison sequences were different. These difference scores provided a measure of the listeners' sensitivity to each of the conditions in which test and comparison sequences were different. Difference scores were subjected to an ANOVA with repeated measures on two factors: attention (2 levels) and condition (switch condition, change in pitch pattern, change in durational pattern, and disjunct condition). Significant effects were confirmed with Greenhouse-Geisser and Huynh-Feldt corrections.

Figure 6 provides mean difference scores for the switch and simple/disjunct conditions. The simple and disjunct conditions were collapsed because mean difference scores for these conditions were not significantly different, either when listeners were attentive [$F(1,14) = 1.52, p = .24$] or when they were distracted [$F(1,14) = 3.18, p = .10$]. Mean difference scores for simple changes in pitch and durational patterns also were not significantly different, either when listeners were attentive [$F(1,14) = .19, p = .67$] or when they were distracted [$F(1,14) = .19, p = .67$].

There was a main effect of attention [$F(1,14) = 14.30, p < .005$], but no main effect of condition [$F(3,14) = 1.49, p = .23$], and there was no significant interaction between attention and condition [$F(3,42) = 1.97, p = .13$]. The three predictions were evaluated in further tests. First, for conditions of attention, the mean difference score for the switch condition was significantly greater than zero [i.e., chance performance;
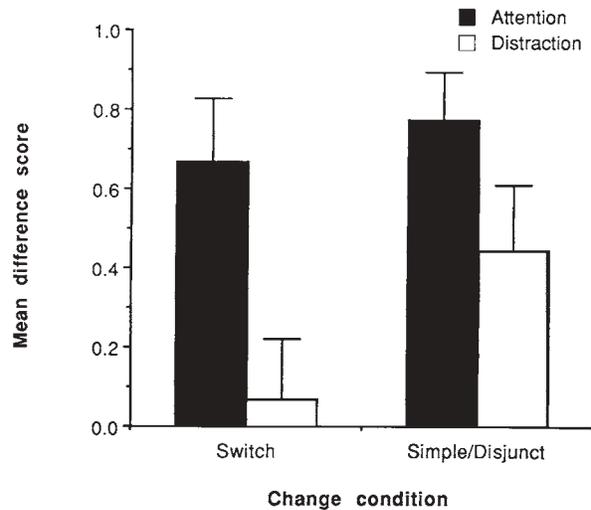

Figure 6. Mean difference scores for the switch condition and the simple-change/disjunct condition, for the attention and distraction conditions of Experiment 4.

$t(14) = 4.18, p < .001$]. This effect confirms that when listeners were attentive, they were sensitive to the manner in which pitch patterns and durational patterns were combined. Second, for conditions of distraction, the mean difference score for the switch condition was not significantly greater than zero [i.e., chance performance; $t(14) = .43, p = .67$]. The poor performance shown by the distracted listeners in the switch condition suggests that attention plays a role in the process of integrating pitch pattern with durational pattern. A comparison of difference scores under the attention and distraction conditions (for switch conditions) confirmed that this influence of attention was reliable [$F(1,14) = 9.95, p < .01$].

Finally, Prediction 3 was partially supported. When the listeners were distracted, they perceived a difference between test and comparison sequences significantly more often for simple-change and disjunct conditions than for switch conditions [$F(1,14) = 7.15, p < .02$]. However, when the listeners were attentive, detection for simple-change and disjunct conditions was not significantly better than detection for switch conditions [$F(1,14) = .46, p = .51$]. That is, the attentive listeners were highly sensitive to the manner in which pitch pattern and durational pattern were combined.

DISCUSSION

The results of this investigation may be summarized as follows. When attentive, listeners were consistently sensitive to the manner in which pitch and duration were combined. However, they were usually less sensitive to combinations of parameters than they were to individual parameters. When distracted, listeners were sensitive to individual parameters, but they usually showed no sensitivity to combinations of parameters.

Different stimulus conditions gave rise to varying degrees of sensitivity to the manner in which pitch and duration were combined. For both individual notes and note sequences, attentive listeners were sensitive to changes in switch conditions, indicating that they were sensitive to pitch/duration combinations. This finding is compatible with the view that some integration of auditory parameters occurs under normal (attentive) listening conditions. Even for attentive listeners, however, this integration process was often only partially successful. In Experiments 1–3, attentive listeners were less sensitive to the manner in which parameters were combined than they were to individual parameters.

A similar difficulty with combinations of musical parameters has been reported by Thompson and Sinclair (1993), Hall and Pastore (1993), and Krumhansl (1991), and it is also compatible with predictions derived from feature-integration theory (Treisman & Gelade, 1980). Effects relying only on sensitivity to individual parameters are predicted to be highly salient, because the analysis of individual parameters is thought to occur automatically and preattentively. Effects relying on sensitivity to combinations of parameters are predicted to be less salient, because the integration of parameters depends more heavily on information-processing conditions (e.g., degree of attention) than does the analysis of individual parameters.

A comparison of the results from Experiments 1–3 suggests that the listeners were least sensitive to combinations of parameters when component pitches were separated by a wide pitch interval and were most sensitive to combinations of parameters when durations were highly distinctive. Experiments 1–3 consistently showed that the listeners were less sensitive to combinations of parameters than to changes in individual parameter values. However, it was demonstrated in Experiment 4 that, by adjusting stimulus conditions, attentive listeners' sensitivity to combinations of parameters can be improved to a level that is similar to that seen for individual parameters. This increase in listeners' ability to encode combinations of parameters was achieved by considering simple scalar pitch patterns and rhythms.

The findings implicate a role for attention in the integration of auditory parameters, which is consistent with feature-integration theory. When the listeners were attentive, they were sensitive to both individual parameters and combinations of parameters. When they were distracted, they were sensitive to individual parameters, but they were usually insensitive to the manner in which the parameters were combined. These findings are compatible with the view that individual parameters are processed preattentively, whereas the integration of parameters is strongly influenced by focused attention. Notably, the results of Experiment 3 also suggest that attention, though sufficient, is not always necessary for listeners to encode combinations of parameters.

Treisman and Gelade's (1980) feature-integration theory provides the most appropriate theoretical basis for interpreting these results, because it describes the processing implications for tasks that assess subjects' ability to detect combinations of stimulus features. However, it may also be informative to discuss the present results in view of auditory research that draws from Garner's (1974) work on dimensional interactions. In a recent set of articles, Melara and Marks (1990a, 1990b, 1990c) provided evidence for an "interactive multichannel" model of auditory processing (see also Krumhansl & Iverson, 1992). In this model, stimulus processing is understood as a bank of higher order channels. Each channel is targeted to a specific primary perceptual dimension, such as frequency, intensity, and duration. However, there may be leakage, or "cross talk," between channels, so that judgments directed toward one dimension may be influenced by variations in irrelevant dimensions. The systematic study of such cross talk has been the focus of Melara and Marks's investigations.

A fundamental difference between the present investigation and those reported by Melara and Marks is that their tasks involved attending to one dimension at a time and examining influences by other dimensions. In the present research, however, listeners' sensitivity to individual parameters was compared with their sensitivity to combinations of parameters. The distracted listeners' ability to perceive simple texture changes confirms that the analysis of primary perceptual dimensions occurs preattentively. Although cross talk between channels for individual parameters was not examined directly in this research, the distracted listeners' insensitivity to combinations of parameters (for most conditions) suggests that little cross talk between channels occurs at a preattentive level of processing.

A comparison of the results for the switch and simple conditions indicates that even attentive listeners are not as sensitive to combinations of parameters as they are to individual parameters. This comparative difficulty in perceiving texture changes for switch conditions is entirely consistent with results reported for visual texture-segregation tasks (Treisman & Gelade, 1980, Experiments 5–7). The difficulty can be interpreted in terms of differences in the accessibility and perceptual salience of preattentive and attentive processes. Without appealing to a preattentive/attentive distinction, the difficulty may be alternatively understood as a kind of cross talk between channels for combinations of parameters (which would code the texture change for switch conditions) and channels for individual parameters (which would not code the texture change). Put more simply, the listeners may have been unable to focus their attention exclusively on the configural aspects of pitch/duration combinations without being influenced by the qualities of the individual parameters.

Regardless of the underlying processes involved, partial sensitivity to combinations of parameters may have important implications for the perception and memory of music. If the goal of a listener is accurate recognition or reproduction of a musical passage, then accurate in-

tegration of musical parameters is necessary. Performers, for example, must encode the precise manner in which musical parameters, such as pitch patterns and durational patterns, are combined. Not all listeners, however, must encode all points of contiguity between durations, pitches, durational patterns, and pitch patterns. It is possible that under many conditions, including states of partial distraction, listeners obtain an appreciation of the "aesthetic ingredients" of music, without maintaining a strict sensitivity to the precise manner in which these qualities are combined (Krumhansl, 1991; Thompson & Sinclair, 1993). As an example, it may be possible to appreciate a rhythmic pattern in music and reproduce that pattern without recalling the particular sequence of pitches associated with that rhythm. As a second example, it may be possible to recall the presence of certain instruments and certain melodies in orchestral music without necessarily recalling which instruments were associated with which melodies.

The present findings extend results reported by Krumhansl (1991). In her study, listeners judged whether two segments of atonal music were from the same piece. Judgments were influenced by individual parameters (e.g., whether the two segments had similar pitch contours) and overall correlations between parameters (whether, in both segments, high frequencies tended to have shorter durations), but they were not influenced by the specific manner in which the musical parameters were combined. Thus, the listeners were either unable to encode the specific combinations of parameters, or they did not consider such combinations to be relevant to their judgments. The present investigation confirms that, at the level of perception and short-term memory, listeners are able to encode combinations of parameters, but they are usually less sensitive to combinations of parameters than they are to individual parameters.

Given that the experience of listening to music involves varying levels of attention and listening conditions, it is likely that there are aesthetic implications of partial sensitivity to the manner in which musical parameters are combined. The compositional technique of introducing variations on a theme, for example, may in some cases be viewed as a dissociation and recombination of musical qualities (for a relevant discussion, see Narmour, 1990, chap. 4). Thus, findings of partial sensitivity to combinations of parameters may help to account for the large range of musical materials that are perceived as variations of an established theme. Partial sensitivity to combinations of parameters also may allow for flexibility of musical expectations, broadening the possibilities of "good continuation."

REFERENCES

Boltz, M., & Jones, M. R. (1986). Does rule recursion make melodies easier to reproduce? If not, what does? *Cognitive Psychology*, **18**, 389-431.

Briand, K. A., & Klein, R. M. (1989). Has feature integration theory come unglued? A reply to Tsal. *Journal of Experimental Psychology: Human Perception & Performance*, **15**, 401-406.

Butler, D. (1979). A further study of melodic channeling. *Perception & Psychophysics*, **25**, 264-268.

Cohen, A., & Ivry, R. (1989). Illusory conjunctions inside and outside the focus of attention. *Journal of Experimental Psychology: Human Perception & Performance*, **15**, 650-663.

Cohen, A. J. (1991). Tonality and perception: Musical scales primed by excerpts from *The Well-Tempered Clavier* of J. S. Bach. *Psychological Research*, **53**, 305-314.

Cuddy, L. L., & Thompson, W. F. (1992). Asymmetry of perceived key movement in chorale sequences: Converging evidence from a probe-tone analysis. *Psychological Research*, **54**, 51-59.

Desain, P. (1992). A (de)composable theory of rhythm perception. *Music Perception*, **9**, 439-454.

Deutsch, D. (1975). Two-channel listening to musical scales. *Journal of the Acoustical Society of America*, **57**, 1156-1160.

Deutsch, D. (1980). The processing of structured and unstructured tonal sequences. *Perception & Psychophysics*, **28**, 381-389.

Fletcher, N. H., & Rossing, T. D. (1991). *The physics of musical instruments*. New York: Springer-Verlag.

Gabrielsson, A. (1973a). Similarity ratings and dimension analyses of auditory rhythm patterns: I. *Scandinavian Journal of Psychology*, **14**, 138-160.

Gabrielsson, A. (1973b). Similarity ratings and dimension analyses of auditory rhythm patterns: II. *Scandinavian Journal of Psychology*, **14**, 161-176.

Garner, W. R. (1974). *The processing of information and structure*. Potomac, MD: Erlbaum.

Garner, W. R., & Felfoldy, G. L. (1970). Integrality of stimulus dimensions in various types of information processing. *Cognitive Psychology*, **1**, 225-241.

Hall, M. D., & Pastore, R. E. (1993, November). *An auditory analogue to feature integration*. Poster presented at the meeting of the Psychonomic Society, Washington, DC.

Ivry, R. B., & Prinzmetal, W. (1991). Effect of feature similarity on illusory conjunctions. *Perception & Psychophysics*, **49**, 105-116.

Jones, M. R. (1987). Dynamic pattern structure in music: Recent theory and research. *Perception & Psychophysics*, **41**, 621-634.

Jones, M. R., Boltz, M., & Kidd, G. (1982). Controlled attending as a function of melodic and temporal context. *Perception & Psychophysics*, **32**, 211-218.

Krumhansl, C. L. (1990). *Cognitive foundations of musical pitch*. New York: Oxford University Press.

Krumhansl, C. L. (1991). Memory for musical surface. *Memory & Cognition*, **19**, 401-411.

Krumhansl, C. L., & Iverson, P. (1992). Perceptual interactions between musical pitch and timbre. *Journal of Experimental Psychology: Human Perception & Performance*, **18**, 739-751.

Livingstone, M. S., & Hubel, D. H. (1987). Psychophysical evidence for separate channels for the perception of form, color, motion and depth. *Journal of Neuroscience*, **7**, 3416-3468.

Longuet-Higgins, H. C., & Lee, C. S. (1982). The perception of musical rhythms. *Perception*, **11**, 115-128.

Melara, R. D., & Marks, L. E. (1990a). Hard and soft interacting dimensions: Differential effects of dual context on classification. *Perception & Psychophysics*, **47**, 307-325.

Melara, R. D., & Marks, L. E. (1990b). Interaction among auditory dimensions: Timbre, pitch, and loudness. *Perception & Psychophysics*, **48**, 169-178.

Melara, R. D., & Marks, L. E. (1990c). Perceptual primacy of dimensions: Support for a model of dimensional interaction. *Journal of Experimental Psychology: Human Perception & Performance*, **16**, 398-414.

Monohan, C. B., & Carterette, E. C. (1985). Pitch and duration as determinants of musical space. *Music Perception*, **3**, 1-32.

Narmour, E. (1990). *The analysis and cognition of basic melodic structures: The implication realization model*. Chicago: University of Chicago Press.

Palmer, C., & Krumhansl, C. L. (1987a). Independent temporal and pitch structures in determination of musical phrases. *Journal of Ex-*

perimental Psychology: Human Perception & Performance, **13**, 116-126.

PALMER, C., & KRUMHANSL, C. L. (1987b). Pitch and temporal contributions to musical phrase perception: Effects of harmony, performance timing, and familiarity. *Perception & Psychophysics*, **41**, 505-518.

PERETZ, I., & KOLINSKY, R. (1993). Boundaries of separability between melody and rhythm in music discrimination: A neuropsychological perspective. *Quarterly Journal of Experimental Psychology*, **46A**, 301-325.

PERETZ, I., & MORAIS, J. (1989). Music and modularity. *Contemporary Music Review*, **4**, 279-293.

THOMPSON, W. F., & MOR, S. (1992). A perceptual investigation of polytonality. *Psychological Research*, **54**, 60-71.

THOMPSON, W. F., & SINCLAIR, D. (1993). Pitch pattern, durational pattern, and timbre: A study of the perceptual integration of auditory qualities. *Psychomusicology*, **12**, 29-47.

TREISMAN, A. (1986). Properties, parts, and objects. In K. R. Boff, L. Kaufman, & J. P. Thomas (Eds.), *Handbook of perception and human performance: Vol. II. Cognitive processes and performance* (pp. 35/1-35/70). Chichester, U.K.: Wiley.

TREISMAN, A. (1991). Search, similarity, and integration of features between and within dimensions. *Journal of Experimental Psychology: Human Perception & Performance*, **17**, 652-676.

TREISMAN, A., & GELADE, G. (1980). A feature-integration theory of attention. *Cognitive Psychology*, **12**, 97-136.

TREISMAN, A., & SATO, S. (1990). Conjunction search revisited. *Journal of Experimental Psychology: Human Perception & Performance*, **16**, 459-478.

TREISMAN, A., & SCHMIDT, N. (1982). Illusory conjunctions in the perception of objects. *Cognitive Psychology*, **14**, 107-141.

NOTES

1. *Sensitivity* is used here to refer to the extent to which listeners can perceive and make reliable judgments about changes in pitch, duration, pitch pattern, and durational pattern.

2. The order in which the three experiments were conducted was randomly and independently determined for each listener.