Journal of Experimental Psychology:
Human Perception and Performance
1992, Vol. 18, No. 3, 739–751

Copyright 1992 by the American Psychological Association, Inc.
0096-1523/92/$3.00

Perceptual Interactions Between Musical Pitch and Timbre

Carol L. Krumhansl and Paul Iverson
Cornell University

Three experiments examined perceptual interactions between musical pitch and timbre. Experiment 1, through the use of the Garner classification tasks, found that pitch and timbre of isolated tones interact. Classification times showed interference from uncorrelated variation in the irrelevant attribute and facilitation from correlated variation; the effects were symmetrical. Experiments 2 and 3 examined how musical pitch and timbre function in longer sequences. In recognition memory tasks, a target tone always appeared in a fixed position in the sequences, and listeners were instructed to attend to either its pitch or its timbre. For successive tones, no interactions between timbre and pitch were found. That is, changing the pitches of context tones did not affect timbre recognition, and vice versa. The tendency to perceive pitch in relation to other context pitches was strong and unaffected by whether timbre was constant or varying. In contrast, the relative perception of timbre was weak and was found only when pitch was constant. These results suggest that timbre is perceived more in absolute than in relative terms. Perceptual implications for creating patterns in music with timbre variations are discussed.

A musical tone can be characterized in terms of four basic psychological attributes: pitch, duration, loudness, and timbre. Pitch and duration are used musically in highly controlled ways to create complex patterns such as melodies, harmonic progressions, meters, and rhythms. These are the primary means through which traditional music is organized. As a consequence, psychological studies of music have tended to focus on pitch and duration (for recent reviews, see Dowling & Harwood, 1986; Handel, 1989; Krumhansl, 1990, 1991). Most studies have treated pitch and durational patterns separately, although a few studies have examined their interactions. Rhythmic structure has been found to affect the accuracy of pitch memory; a pitch that is accented or rhythmically differentiated from its context is remembered better than one that is not (M. R. Jones, Kidd, & Wetzel, 1981; M. R. Jones, Boltz, & Kidd, 1982). However, when temporal and pitch groupings do not correspond, memory is impaired (Boltz & M. R. Jones, 1986; Deutsch, 1980; Monahan, Kendall, & Carterette, 1982; but see K. C. Smith & Cuddy, 1989). Finally, M. R. Jones, Summerell, and Marshburn (1987) found that presenting a melody in a different rhythm from the original rhythm made it difficult to recognize. These results suggest interactions between pitch and duration in music perception and memory. Although both pitch and durational patterns have been shown to contribute to perceived phrase structure (Palmer & Krumhansl, 1987a, 1987b) and melodic similarity (Monahan & Carterette, 1985), their contribution to these perceptual attributes appears to be largely independent. Thus, interactions between these dimensions may depend on the perceptual attribute that is being measured.

Variations in loudness (dynamics) are used in music to highlight structural units. For example, loudness contrasts may signal phase boundaries, aid in establishing a regular metrical framework, or emphasize significant melodic, harmonic, or rhythmic events. This aspect of music has received relatively little experimental analysis. In a recent study, however, Nakamura (1987) demonstrated that dynamic markings in music were communicated quite well to listeners. This was true even though physical intensity did not correspond well to either the notated or perceived dynamics, which suggests that context influences the way loudness variations are performed and interpreted. Moreover, perceived dynamics were influenced by other musical attributes. For example, rising pitch enhanced the impression of crescendos (increasing loudness), whereas falling pitch enhanced the impression of decrescendos (decreasing loudness). Thus, it appears that loudness interacts in perception with other musical attributes.

In the present study, we examine whether timbre also interacts with other attributes of musical tones. In particular, we focus on interactions between timbre and pitch. Timbre, sometimes referred to as *tone quality* or *tone color*, is defined as the way in which musical sounds differ once they are equated for pitch, loudness, and duration. For example, it is how a violin and an oboe differ when they are playing the same pitch for the same duration at the same loudness. Contemporary composers and music theorists (e.g., Cogan, 1984; Erickson, 1975; Lerdahl, 1987; Slawson, 1985) have shown an increased interest in understanding how timbre is, or might be, used compositionally. It is possible to construct patterns of timbres that have properties analogous to patterns of pitch and duration? For example, are timbres heard as if they are hierarchically organized (Lerdahl, 1987), and are sequences of timbres subject to such transformations as transposition and inversion (Slawson, 1985)? These questions highlight the fact that we know relatively little about timbre from

---

This research was supported by a grant to Paul Iverson from Sigma Xi, The Scientific Research Society.

We are grateful to John Pinto and Karen Jacowitz for their assistance in conducting the experimental sessions and to Robert Crowder, Robert Melara, and two anonymous reviewers for their detailed and very constructive suggestions for revising an earlier version of this article.

Correspondence concerning this article should be addressed to Carol L. Krumhansl, Department of Psychology, Cornell University, Ithaca, New York 14853.

a psychological point of view. The experiments reported here investigated whether pitch and timbre are coded independently in perception, both for isolated tones and for tones presented in sequences. Even if timbre can be manipulated independently from the other attributes of musical sound, as is assumed by its definition, does it follow that timbre functions independently of the other attributes in perception and memory?

To date, the primary issue motivating perceptual studies of timbre has been the relationship between acoustic and perceptual properties. Acoustic analyses show that musical instruments have complex characteristics (for reviews, see Benade, 1976; Risset & Wessel, 1982). During the steady-state portion, the relative amplitudes of the different harmonics vary from instrument to instrument. Moreover, the amplitudes of the harmonics change over time, and the pattern of these amplitude functions, especially during onset and offset portions, are quite distinctive for different instruments. A variety of methods have been used to isolate the acoustic properties that determine timbre perception, including multidimensional scaling (Grey, 1977; Grey & Gordon, 1978; Miller & Carterette, 1975; Wessel, 1973), factor analysis (Wedin & Goude, 1972), adjective ratings (von Bismarck, 1974), and instrument identification (Berger, 1964; Grey, 1978; Kendall, 1986; Saldanha & Corso, 1964; Wedin & Goude, 1972). These investigations show that timbre is a multidimensional attribute. Some of the dimensions that are perceptually important include the brightness of the sound (the relative amplitude of the higher harmonics), the rapidity of the attack portion, whether the rise of the different harmonics is synchronous, the presence or absence of noise during the onset, and the presence or absence of even harmonics. In addition to these common dimensions, a scaling study done by Wessel, Winsberg, and Krumhansl (reported in Krumhansl, 1989) showed that some instrument timbres may have distinctive characteristics that set them apart from all other instruments. For the most part, these studies used isolated tones matched for pitch, loudness, and duration, in keeping with the definition of timbre. As such, they do not assess possible interactions between timbre and the other attributes of musical tones.

A few recent studies suggest that the timbre and pitch of a tone are not perceived independently. In one experiment, Beal (1985) asked listeners to judge whether two chords played on guitar, piano, and harpsichord were the same or different (did or did not contain the same pitches). Nonmusicians found it very difficult (error rate of 64%) to judge two chords as the same when they were played on different instruments, even though they made few errors when the chords were played on the same instrument. Musicians were more accurate overall but still found it more difficult to compare chords when they were played on different instruments than when they were played on the same instrument. In a second experiment, Beal asked listeners to judge whether two chords were played by the same instrument. Nonmusicians found it more difficult to judge the instruments as the same when the chords were different than when they were the same. Musicians were quite accurate overall but made slightly more errors for different chords than for the same chords. These findings suggest that it is difficult to filter out variations in timbre when judging pitch and variations in pitch when judging timbre, especially for nonmusicians. Crowder (1989) obtained similar effects of timbre on pitch judgments. Same pitches were judged faster and more accurately when they were played on the same instrument (guitar, flute, or trumpet) than when they were played on different instruments. This effect was found even when the first of the comparison tones was imagined by the subject. Written instructions indicated the instrument they were to imagine playing the tone, and a neutral sine wave indicated its pitch. Crowder did not investigate either the effect of pitch differences on timbre judgments or systematic differences between musicians and nonmusicians.

In an extensive series of experiments, Melara and Marks (1990a, 1990b, 1990c) investigated the perceptual interactions among timbre, pitch, and loudness. With sorting tasks in the tradition of Garner (1974, 1981; Pomerantz & Pristach, 1989), subjects were instructed to classify individual tones as quickly as possible according to a given classification scheme that varied between blocks of trials. The interaction between stimulus dimensions was assessed by the pattern of reaction time differences across the classification tasks. The standard Garner sorting tasks, which are used in our first experiment, are described in more detail later. Briefly, however, dimensions are said to interact if when sorting on the basis of one relevant dimension, uncorrelated variation on the irrelevant dimension interferes with performance, and correlated variation on the irrelevant dimension facilitates performance. This is the pattern Melara and Marks (1990b) found for timbre and pitch, leading them to conclude that these dimensions interact in individual tones, which is consistent with the studies just cited. They also found that timbre and loudness interact.

By extending the standard sorting tasks, Melara and Marks (1990a, 1990c) uncovered additional information about the perceptual processing of timbre, pitch, and loudness. With their "multiclass" procedure (Melara & Marks, 1990a), they examined the ability of listeners to maximize performance on classification tasks by using the facilitation of correlated variation between dimensions while ignoring the interference of uncorrelated variation. For responses that were based on timbre or pitch, listeners were able to ignore uncorrelated variation while using correlated variation. In additional studies, Melara and Marks (1990c) examined how different combinations of timbre, pitch, and loudness influence the use of correlated and uncorrelated variation. They found that listeners were best able to maximize the effect of correlated variation and minimize the effect of uncorrelated variation when there was uncorrelated variation on only a single attribute (orthogonal combinations of values) rather than on two attributes (nonorthogonal combinations of values). For present purposes, the important conclusion of the Melara and Marks (1990a, 1990b, 1990c) studies is that classification tasks consistently demonstrate interactions between timbre and both pitch and loudness. All of these studies, however, used nonmusical timbres. Different timbres were produced by changing the width-to-cycle-length ratio of a rectangular wave form. This produces tones varying along a continuum described as changing from "twangy" to "hollow." It is of interest, therefore, to see if their basic finding of timbre–pitch interactions is replicated by using more realistic musical timbres, which was the focus of our first experiment.

Thus, experimental evidence increasingly suggests that the timbre and pitch of a single tone are not perceptually independent. This nonindependence has implications for the definition of timbre and its study from a psychoacoustic point of view. Considered from a musical point of view, timbre raises additional psychological questions. How do timbre and pitch function in longer sequences? Does the pitch pattern of a sequence of tones affect the perception of timbre, and does the timbre pattern of a sequence of tones affect the perception of pitch? If so, then the formation of timbre patterns will be affected by pitch variations, and vice versa. Little evidence exists on this question, but two studies suggest that pitch and timbre compete in the formation of auditory "streams" (Bregman & Campbell, 1971), the segregation of a sequence of tones into separate groups. Studies by J. Smith, Hausfeld, Power, and Gorta (1982) and Singh (1987) found that tone sequences can be segregated either on the basis of pitch similarity or on the basis of timbre similarity and that there are trade-offs between these two attributes. It appears, then, that both attributes contribute to the perceptual organization of tone sequences and that they can be set in opposition. A recent study of pitch memory by Semal and Demany (1991), however, finds independence of pitch and timbre. They used the pitch-recognition paradigm of Deutsch (1972) in which two tones (called *standard* and *comparison* tones) are judged as same or different. A number of intervening tones were presented between the standard and comparison tones. Consistent with Deutsch's (1972) results, Semal and Demany (1991) found that intervening tones close in pitch to the standard and comparison tones produced memory interference. They found no effect, however, of the timbre similarity of the intervening tones to the standard and comparison tones, suggesting that pitch memory is independent of the timbre of context tones. They did not consider pitch effects on timbre memory. Thus, as with pitch and duration, the independence of pitch and timbre appears to depend on the psychological attribute that is being measured.

The second and third experiments reported here took a similar approach to that of Semal and Demany (1991) to the issue of how the timbre and pitch of successive tones interact. A target (to-be-remembered) tone was embedded in a sequence, and listeners were instructed to attend to either its pitch or its timbre. We evaluated the effect on memory of changing the tones in the context (the tones preceding and following the target tone). Does changing the pitch of surrounding tones interfere with timbre memory? Does changing the timbre of surrounding tones interfere with pitch memory? If interference is found, then it implies that one attribute's perception is unstable in the sense that it is influenced by the other attribute of context tones. If no interference is found, then this allows for the possibility that patterns formed of one attribute might be perceived as invariant despite variations on the other attribute. Note that the interaction between timbre and pitch of successive tones is logically independent of the interaction of timbre and pitch of a single tone, as was studied in our first experiment and related studies summarized earlier.

The final issue addressed here is the way in which timbres in sequences are encoded with respect to one another. Are timbre values perceived in relation to other timbre values in a way that is analogous to relative pitch perception? This question arises from the analogy with musical pitch. Numerous observations support the idea that a musical pitch is coded primarily in terms of the intervals it forms with simultaneous and successive pitches. For example, a transposed melody is heard as the same melody even though it is shifted up or down in pitch because the relations between tones (the intervals) are unchanged. Do timbre intervals exist, making possible transposition and other transformations as has been suggested by Slawson (1985)? The study by Ehresman and Wessel (1978) provides the only perceptual evidence addressing this question. Their study used 15 synthesized, naturalistic instrument tones of different timbres (all of the same pitch). Multidimensional scaling of these tones produced a two-dimensional representation that was then used to construct timbre analogy problems of the following form: A is to B as C is to D. For each triple of timbres, A, B, C, four different alternatives were provided: D_1, D_2, D_3, and D_4. Listeners most frequently chose the alternative that was closest to the ideal point in the two-dimensional representation defined as the point that completes a parallelogram with the other three timbres. In other words, they chose the alternative D_i that was (as much as possible) the same distance and direction from C as B is from A. It is as though they were able to transpose the timbre interval from A to B to match the interval from C to D_i. In the present article, we take a somewhat different approach to studying whether timbre is perceived relationally. A target tone is embedded in a sequence, and the effect of changing the timbre of contextual tones on memory for the timbre of the target tone is measured.

The various forms of interactions between pitch and timbre that have been discussed can be schematically summarized by the diagram in Figure 1. The sequence in the example consists of seven tones, each of which has a particular value of pitch (P) and a particular value of timbre (T), which may vary from tone to tone. If the pitch and timbre of individual tones are not processed independently, this is indicated by vertical lines showing that these two attributes interact. If pitches of successive tones are not processed independently, this is indicated by horizontal lines connecting the different pitches. Similarly, if timbres of successive tones are not processed independently, this is indicated by horizontal lines connecting the different timbres. Finally, if pitch is not proc-

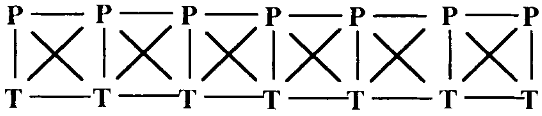

Figure 1. The kinds of possible perceptual interactions between musical pitch and timbre. (The diagram schematically illustrates a sequence of seven tones. Vertical lines indicate that the timbre and pitch of single tones are not independent. Horizontal lines indicate the nonindependence of the same attribute [pitch or timbre] of successive tones. Diagonal lines indicate the nonindependence of different attributes [pitch and timbre] of successive tones.)

essed independently of the timbres of surrounding tones, or if a timbre is not processed independently of the pitches of the surrounding tones, this is indicated by the diagonal lines. Evidence concerning these various kinds of perceptual interaction comes from assessing listeners' ability to attend to one attribute (pitch or timbre) of a target tone presented either in isolation (Experiment 1) or in sequences (Experiments 2 and 3).

Experiment 1

The first experiment used the standard Garner (1974) sorting tasks to evaluate how pitch and timbre interact at the level of individual tones. In this methodology, subjects are asked to categorize stimuli into two groups as fast as possible according to a specified criterion. Reaction times and errors are recorded. Figure 2 schematically describes the stimuli. The stimuli vary on two attributes (pitch and timbre) with two levels for each attribute, giving four possible attribute combinations, denoted P_1T_1, P_1T_2, P_2T_1, and P_2T_2. Table 1 shows the tasks. In control tasks, subjects are asked to categorize two stimuli that vary on a single attribute (e.g., P_1T_1 vs. P_1T_2). In correlated-dimension tasks, subjects are asked to categorize two stimuli that are different on both attributes (e.g., P_1T_1 vs. P_2T_2). In selective-attention (or filtering) tasks, subjects are asked to categorize the four stimuli on the basis of a single attribute while ignoring the uncorrelated variation in the other attribute (e.g., P_1T_1 and P_2T_1 vs. P_1T_2 and P_2T_2). In the condensation (divided-attention) task, subjects are asked to categorize the four stimuli into two groups that are not defined by a single attribute (P_1T_1 and P_2T_2 vs. P_1T_2 and P_2T_1). In focusing tasks, subjects are asked to categorize one tone into one group and the other three tones into the other group (e.g., P_1T_1 vs. P_1T_2 and P_2T_1 and P_2T_2).

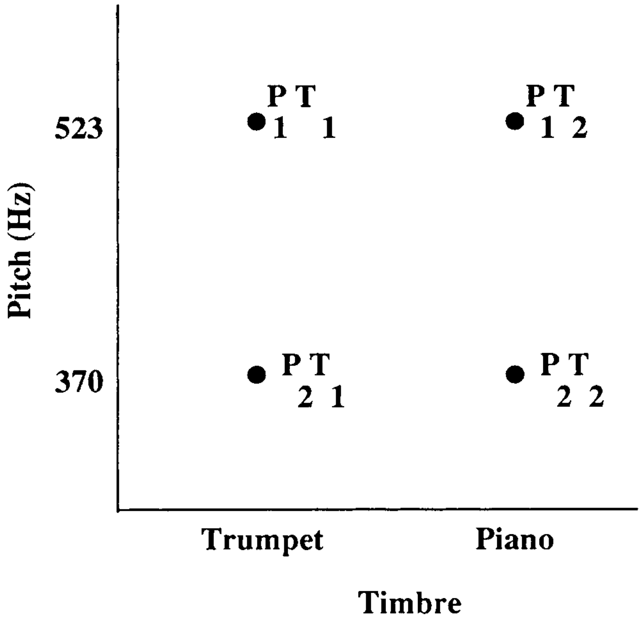

Figure 2. The combinations of pitch (P) and timbre (T) values used in Experiment 1, which used the standard Garner sorting tasks described in Table 1.

Comparing the mean reaction times for these tasks yields information about the perceptual independence of the attributes of pitch and timbre. If the uncorrelated variation of the selective-attention (filtering) tasks does not interfere with categorization when compared with the control tasks, then selective attention is possible. If there is facilitation from the correlated variation in the correlated-dimension tasks when compared with the control tasks, then either categorization is made on the easiest-to-categorize attribute or both attributes are attended simultaneously (Pomerantz & Pristach, 1989). If the condensation task takes longer than both the control tasks and the selective-attention (filtering) tasks, then divided attention is difficult. Finally, differences between the focusing tasks examine configural properties of the stimuli; if one of the focusing tasks is easier than the others, then a particular combination of attributes is easier to classify. With this information, the interaction between attributes can be evaluated in terms of Garner's (1974) classification scheme. Attributes are said to be *Garner separable* if selective attention is possible, there is no facilitation in the correlated-dimension tasks, and divided attention is difficult. Attributes are said to be *Garner integral* if selective attention is not possible, there is facilitation in the correlated-dimension task, and divided attention is difficult. Finally, attributes are said to be *Garner configural* if selective attention is not possible, there is no facilitation in the correlated-discrimination task, and divided attention is as easy as the other tasks.

This experiment essentially replicates that reported by Melara and Marks (1990b). The primary difference was that here we used naturalistic musical timbres. Figure 3 illustrates some of the complex acoustic properties that distinguish between musical timbres. The top left panel shows the amplitude

Table 1
Experiment 1: Categorization Tasks and Results

Categorization task	Task type	Reaction time (ms)
P_1T_1 versus P_1T_2	Control	361
P_2T_1 versus P_2T_2	Control	362
P_1T_1 versus P_2T_1	Control	320
P_1T_2 versus P_2T_2	Control	317
P_1T_1 versus P_2T_2	Correlated dimension	304
P_1T_2 versus P_2T_1	Correlated dimension	307
P_1T_1 and P_2T_1 versus P_1T_2 and P_2T_2	Selective attention	420
P_1T_1 and P_1T_2 versus P_2T_1 and P_2T_2	Selective attention	405
P_1T_1 and P_2T_2 versus P_1T_2 and P_2T_1	Condensation	708
P_1T_1 versus P_1T_2 and P_2T_1 and P_2T_2	Focusing	434
P_1T_2 versus P_2T_1 and P_2T_2 and P_1T_1	Focusing	447
P_2T_1 versus P_2T_2 and P_1T_1 and P_1T_2	Focusing	418
P_2T_2 versus P_1T_1 and P_1T_2 and P_2T_1	Focusing	472

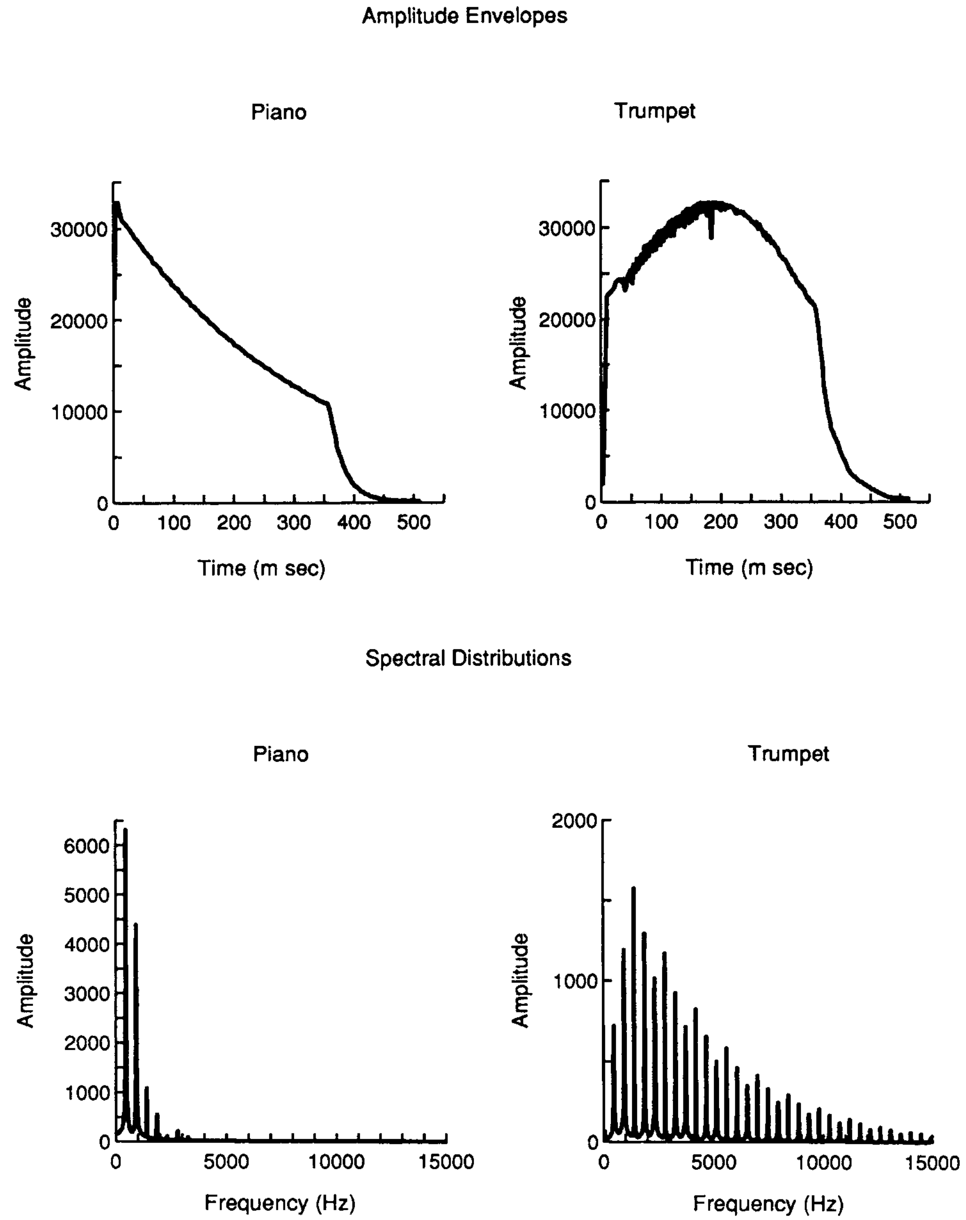

Figure 3. The acoustic differences between the piano and trumpet tones used in Experiment 1. (The top panels show the amplitude envelopes, which plot the amplitude [on an arbitrary linear scale] as a function of time. The piano timbre is characterized by a drop in amplitude after the initial rise, whereas the trumpet tone has a sustained high amplitude after the initial rise. The bottom panels show the spectral distributions, which plot the amplitude as a function of frequency. The piano timbre has relatively strong low harmonics, whereas the trumpet tone has relatively strong high harmonics.)

envelope of the piano timbre used in the experiment, that is, how the amplitude changes over time. (Amplitude is measured on an arbitrary linear scale.) Following a sharp rise, the amplitude decays gradually over the rest of the tone, producing a percussive sound. In contrast, the trumpet's amplitude envelope has a gradual onset and then stays at a relatively high level before a rapid decay. The bottom left panel shows the spectral distribution of the steady-state portion of the piano timbre. The function plots the amount of energy at each frequency. As is apparent, energy is present primarily at a few low-frequency harmonics. The bottom right panel shows the spectral distribution of the trumpet timbre; it contains energy at many higher frequency harmonics, producing a bright sound. These two timbres, which are easily discriminable, were chosen to represent the range of variation exhibited by musical timbres found in scaling experiments of orchestral instruments (e.g., Iverson & Krumhansl, 1991; Krumhansl, 1989). The musical pitches used in the experiment were separated by an interval of a tritone (a half octave). They were also easily discriminable and have been found to be maximally separated in scaling studies of tones in an octave range (Krumhansl, 1979, 1990).

Method

Subjects. Subjects were 11 members of the Cornell University community who were each paid for participating in the half-hour experiment. Each subject had studied at least one musical instrument for a minimum of 5 years ending not more than 3 years before participating in this study. One additional subject's data file was lost because of computer error.

Apparatus. Stimuli were generated by a Yamaha TX-816 frequency modulation tone generator that was controlled by an IBM-AT microcomputer through a Roland MPU 401 MIDI interface. Stimuli were amplified and presented to each subject over a single high-quality speaker in a soundproof room. Subject responses were entered and recorded by the microcomputer, which controlled the presentation of stimuli.

Stimulus materials. The four tones used in the experiment were generated by varying two levels of pitch and two levels of timbre. The pitch of each tone was either 370 or 523 Hz, corresponding to the musical tones F#4 and C5, respectively (where 4 denotes the octave beginning with middle C). The timbre of each tone was either a synthesized trumpet or a synthesized piano. Each tone was presented at 55 dBA SPL and was sounded until the subject responded.

Procedure. Subjects read that they would hear tones that they would be asked to categorize in several different ways. Before each block, they would hear over the speaker and read on the computer screen which tones belonged in which category (e.g., Category 1 = high trumpet; Category 2 = high piano). They were told to press either the *1* or *2* key on the computer keyboard to indicate the category. They were asked to respond as quickly as possible without making mistakes but were told that speed was more important than accuracy. After the instructions, subjects heard the four tones repeatedly until they were familiar with them. After this, they completed two practice blocks that were eight trials long while the experimenter was present to answer questions and monitor performance. One of these practice blocks was a randomly selected correlated-dimension task, and the other was a randomly selected selective-attention task. After these, the experimenter left the room, and each subject completed the 13 blocks of experimental trials in an order chosen randomly for each subject. The 13 blocks of experimental trials correspond to the categorization tasks shown in Table 1. The first 4 trials in a block were practice, and the next 36 were experimental. Listeners were permitted to take short breaks between blocks.

Results and Discussion

Table 1 shows the mean reaction time for correct responses in each condition. Error rates were too low to be useful indicators of performance. They correlated positively with reaction times, indicating no speed–accuracy trade-off, $r(11) = .718$, $p < .01$. The main analysis was a comparison of the reaction times on trials with correct responses in the first nine conditions. The focusing tasks did not differ from one another, $F(3, 30) = 2.772$, $p > .05$, indicating that the various timbre–pitch combinations were approximately equally easy to categorize. The analysis of the nine conditions found a significant effect of task, $F(8, 80) = 65.403$, $p < .001$. We performed *t* tests to assess the pattern of differences between task means, which was as follows. The mean (340 ms) of the four control tasks (P_1T_1 vs. P_1T_2, P_2T_1 vs. P_2T_2, P_1T_1 vs. P_2T_1, and P_1T_2 vs. P_2T_2) was significantly greater than the mean (306 ms) of the two correlated-dimension tasks (P_1T_1 vs. P_2T_2; P_1T_2 vs. P_2T_1), $t(10) = 5.323$, $p < .001$, indicating facilitation from correlated variation in pitch and timbre.

The mean (362 ms) of the two control tasks in which timbre defines the categories (P_1T_1 vs. P_1T_2; P_2T_1 vs. P_2T_2) was significantly less than the mean (420 ms) of the selective-attention (filtering) task in which timbre defines the categories (P_1T_1 and P_2T_1 vs. P_1T_2 and P_2T_2), $t(10) = -4.216$, $p < .01$, indicating interference from uncorrelated variation in pitch. Similarly, the mean (319 ms) of the two control tasks in which pitch defines the categories (P_1T_1 vs. P_2T_1; P_1T_2 vs. P_2T_2) was significantly less than the mean (405 ms) of the selective-attention (filtering) task in which pitch defines the categories (P_1T_1 and P_1T_2 vs. P_2T_1 and P_2T_2), $t(10) = -4.271$, $p < .01$, indicating interference from uncorrelated variation in timbre. Selective attention was symmetrically difficult.

The mean (413 ms) of the two selective-attention (filtering) tasks (P_1T_1 and P_2T_1 vs. P_1T_2 and P_2T_2; P_1T_1 and P_1T_2 vs. P_2T_1 and P_2T_2) was significantly less than the mean (708 ms) of the condensation task (P_1T_1 and P_2T_2 vs. P_1T_2 and P_2T_1), $t(10) = -8.419$, $p < .001$. In addition, the mean (340 ms) of the four control tasks (P_1T_1 vs. P_1T_2, P_2T_1 vs. P_2T_2, P_1T_1 vs. P_2T_1, and P_1T_2 vs. P_2T_2) was significantly less than the mean (708 ms) of the condensation task (P_1T_1 and P_2T_2 vs. P_1T_2 and P_2T_1), $t(10) = -10.603$, $p < .001$. Thus, divided attention was found to be difficult by either contrast.

These results confirmed the conclusion of Melara and Marks (1990b) that pitch and timbre are Garner integral. Although the magnitude of the differences between tasks was greater in our study, the general pattern of results was the same. Apparently, subjects could not attend to the pitch of a tone without being influenced by its timbre and could not attend to the pitch of a tone without being influenced by its pitch. (This kind of interaction is represented by the vertical lines of Figure 1.) The interaction effects were symmetrical. This was true even though the baseline discrimination tasks for pitch were somewhat faster than those for timbre. This reaction time difference is hard to interpret, however, given that some of the critical information for timbre, the amplitude envelope, evolves over time. Despite these interaction effects, the low error rates show that listeners could recognize that two tones had the same pitch even if they were played on different instruments, and they could recognize that two tones were played on the same instrument even if they had different pitches. Thus, although pitch and timbre of a tone interact in the sense of being Garner integral, pitch and timbre values can be abstracted from the stimulus tones, allowing such classifications to be made accurately. This is consistent with Melara and Marks's (1990c) claim that timbre and pitch are primary perceptual attributes in auditory perception. It is clear from these results that pitch and timbre interact at the level of individual tones. Do these attributes also interact in sequences of tones? The perception of timbre and pitch of successive tones may be subject to different processes than isolated tones, a possibility explored in the next two experiments.

Experiment 2

This experiment investigated possible pitch and timbre interactions between successive tones in sequences. The evidence for interactions comes from tasks in which listeners were instructed to attend to one of the two attributes. Thus, these are like the selective-attention tasks of the first experiment in the sense that only one aspect of the stimulus is relevant to the required response. Here, however, interference from the tones in the context (tones preceding and following a designated target tone) was assessed rather than the variation of the other attribute of the test tone. On some trials, pitch was the relevant attribute; on others, timbre was the relevant attribute. Listeners were required to make same–different judgments of the target tone (always in one designated position) in two sequences of tones presented in succession. The context was either the same, different on the relevant attribute, or different on the irrelevant attribute. The irrelevant attribute of the target tone was never changed.

More specifically, subjects listened to pairs of seven-tone sequences, the first of which is called the standard sequence, the second of which is called the comparison sequence. They were instructed to attend to changes on a single attribute of the fourth tone, which was the designated target. Figure 4 illustrates the conditions of the experiment through sample trials on which pitch was the relevant attribute; the trials on which timbre was the relevant attribute were constructed analogously. All the examples shown are different trials, as is indicated by the change in pitch between the fourth serial position of the first, standard sequence and the fourth serial position of the second, comparison sequence. On trials in the no-context-change condition, the corresponding context tones were the same in terms of both pitch and timbre. (Within a sequence both pitch and timbre varied; the values were randomly selected from a set of eight pitches and eight timbres.) On trials in the relevant-attribute context-change condition, corresponding context tones were the same on the irrelevant attribute, and corresponding context tones were different on the relevant attribute. Thus, for pitch-relevant trials there were changes in the pitches of the context tones but no change of their timbres. On trials in the irrelevant-attribute context-change condition, corresponding context tones were the same on the relevant attribute, and corresponding context tones were different on the irrelevant attribute. Thus, for pitch-relevant trials there were changes of the timbres of the context tones but no changes of their pitches.

Relevant Attribute: Pitch

	Standard	**Comparison**
No Context Change	$P_1 P_8 P_4 \mathbf{P_3} P_2 P_7 P_6$ $T_3 T_4 T_6 \mathbf{T_2} T_8 T_5 T_1$	$P_1 P_8 P_4 \mathbf{P_5} P_2 P_7 P_6$ $T_3 T_4 T_6 \mathbf{T_2} T_8 T_5 T_1$
Relevant-attribute Context Change	$P_1 P_8 P_4 \mathbf{P_3} P_2 P_7 P_6$ $T_3 T_4 T_6 \mathbf{T_2} T_8 T_5 T_1$	$P_8 P_6 P_1 \mathbf{P_5} P_4 P_7 P_2$ $T_3 T_4 T_6 \mathbf{T_2} T_8 T_5 T_1$
Irrelevant-attribute Context Change	$P_1 P_8 P_4 \mathbf{P_3} P_2 P_7 P_6$ $T_3 T_4 T_6 \mathbf{T_2} T_8 T_5 T_1$	$P_1 P_8 P_4 \mathbf{P_5} P_2 P_7 P_6$ $T_1 T_5 T_6 \mathbf{T_2} T_3 T_8 T_7$

Figure 4. Example sequences illustrating the pitch-relevant conditions of Experiment 2; timbre-relevant conditions were constructed analogously. (The target tone was in the fourth serial position. In the no-context-change condition, both pitch and timbre contexts [the tones preceding and following the target] were the same in standard and comparison sequences. In the relevant-attribute context-change condition, the pitch context was changed. In the irrelevant-attribute context-change condition, the timbre context was changed. All of the sample trials illustrated are different trials.)

Two kinds of independence are possible, which we refer to as *same-attribute independence* and *different-attribute independence.* Same-attribute independence refers to the independence between successive tones on the same attribute, that is, the independence of the pitch of a tone from the pitches of the context tones, or the independence of the timbre of a tone from the timbres of the context tones. (The horizontal lines of Figure 1 indicate a lack of same-attribute independence.) Different-attribute independence refers to the independence between tones on different attributes, that is, the independence of the pitch of a tone from the timbres of the context tones and the independence of the timbre of a tone from the pitches of the context tones. (The diagonal lines of Figure 1 indicate a lack of different-attribute independence.) Both kinds of independence are possible. If there is no interference in the relevant-attribute context-change condition (no decrease in performance compared with the no-context-change condition), then there is same-attribute independence. If there is no interference in the irrelevant-attribute context-change condition (compared to the no-context-change condition), then there is different-attribute independence.

The set of eight timbres used in this and the following study was chosen to represent the distribution of timbres found in scaling studies of orchestral instruments (e.g., Iverson & Krumhansl, 1991; Krumhansl, 1989). Because we lacked information about principles governing the formation of timbre patterns, no attempt was made to use systematic rules to construct the timbre sequences. The sequential ordering of timbres was determined by a random selection (without replacement) from the set of eight timbres. In a similar spirit, the eight pitches were chosen to span a musical octave. The octave was selected to be within the natural range of multiple instruments. The pitches form what is known as the octatonic scale. This is a rare musical scale, which does not give a strong impression of tonality (major or minor key; see Krumhansl & Schmuckler, 1986, for more information on this pitch set and its perceptual effects). To further avoid familiar melodic patterns, the pitch values were selected randomly (without replacement) from the set. Thus, any differences between pitch and timbre produced by the experimental manipulations cannot be attributed to differential familiarity of the patterns formed by pitch and timbre sequences.

Method

Subjects. Subjects were 22 members of the Cornell University community who were paid for participating in the 1-hr experiment.

They were selected according to the same criteria as Experiment 1. The data of 4 of the subjects were dropped from the analyses because they performed at chance.

Apparatus. The apparatus was the same as in Experiment 1.

Stimulus materials. The tones used in the stimulus sequences varied in pitch and timbre. Possible pitch values came from the following set: 370, 392, 440, 466, 523, 554, 622, and 659 Hz, which correspond to the musical pitches F#4, G4, A4, A#4, C5, C#5, D#5, and E5, respectively. The timbre of the tones was synthesized to simulate a trumpet, harpsichord, piano, guitar, harp, flute, clarinet, or an oboe. The piano and trumpet timbres were the same as those in Experiment 1. The interval between tone onsets was 360 ms. Each tone ended approximately 40 ms before the onset of the next tone, varying somewhat with timbre. All tones were presented at 55 dBA SPL.

Each trial consisted of a pair of seven-tone sequences, and the target tone was always in the fourth serial position. For the standard sequence, the pitch and timbre values were selected randomly without replacement. The target tone in the comparison sequence was selected as follows. On half of the trials (same trials), it was identical to the target in the standard sequence on both the relevant and the irrelevant attribute. On half of the trials (different trials), it was changed to a different value on the relevant attribute (one not occurring elsewhere in the sequence), and its value on the irrelevant attribute was unchanged. The three conditions (no context change, relevant-attribute context change, and irrelevant-attribute context change) for the pitch-relevant and timbre-relevant trials were constructed as described earlier.

Procedure. Subjects read that they would hear pairs of seven-tone melodies and that they should try to detect a change on the fourth tone (the target tone) in either pitch or timbre (specified for each block of trials) while still listening to the other tones. They were instructed that on each trial, they would hear a pair of sequences and then would enter an integer response from 1 (very sure no change) to 6 (very sure change). They read that the trials were organized into four experimental blocks with the same relevant attribute within each block.

After the instructions, each subject was presented a training block of 12 practice trials with feedback (two trials for each condition for each relevant attribute). The experimenter was present for this block to answer questions and to see if each subject understood the instructions correctly. After this was finished, the experimenter left the room, and each subject completed the four blocks. Subjects were free to take short breaks between blocks.

Each subject was presented two blocks of trials for each of the two relevant attributes (pitch and timbre) in an ordering counterbalanced between subjects. Each block started with 3 practice trials with feedback, continued with 3 practice trials without feedback, and concluded with 48 experimental trials without feedback. The experimental trials were a random ordering of 16 each (8 same and 8 different) for the three conditions (no context change, relevant-attribute context change, and irrelevant-attribute context change).

Results and Discussion

We calculated individual memory-operating characteristics (MOCs) (Swets, 1973) for each subject for each condition. The area under the MOC curve was used as the dependent measure in the analyses. This measure is monotonically related to a d' calculated with signal detection theory but does not assume normal distributions. On this measure, chance performance is .50. Table 2 presents the results for both the pitch-relevant and timbre-relevant conditions. The MOC curves were based on ratings for 16 same and 16 different trials.

Table 2
Experiment 2: Conditions and Results

Condition	Area under memory-operating characteristic curve
Relevant attribute: Pitch	
No context change	.857
Relevant-attribute context change	.807
Irrelevant-attribute context change	.872
Relevant attribute: Timbre	
No context change	.889
Relevant-attribute context change	.909
Irrelevant-attribute context change	.876

For the pitch-relevant conditions, there was a significant overall effect of condition, $F(2, 34) = 4.256$, $p < .05$. Each of the relevant-attribute and irrelevant-attribute context-change conditions was compared with the no-context-change condition by using t tests with the following results. The mean of the relevant-attribute context change condition (.807) was marginally lower than the mean of the no-context-change condition (.857), $t(17) = -2.023$, $p = .059$, indicating interference from changes in pitch context. Thus, the target pitch appears to be coded in relation to other pitches in its context even though the task allowed listeners to focus on a fixed target position. The mean of the irrelevant-attribute context-change condition (.872) was not statistically different from the mean of the no-context-change condition (.857), $t(17) = 0.609$, $p > .05$, indicating no consistent interference from changes in timbre context. Thus, the target pitch appears to be coded independently of its timbre context.

For the timbre-relevant conditions, there was no significant overall effect of condition, $F(2, 34) = 2.091$, $p > .05$. Each of the relevant-attribute and irrelevant-attribute context-change conditions was compared with the no-context-change condition with the following results. The mean of the relevant-attribute context-change condition (.909) was not significantly different from the mean of the no-context-change condition (.889), $t(17) = 1.251$, $p > .05$, indicating no consistent interference from changes in the timbre context. The mean of the irrelevant-attribute context-change condition (.876) was not significantly different from the mean of the no-context-change condition (.889), $t(17) = -0.713$, $p > .05$, indicating no consistent interference from changes in the pitch context. Thus, the target timbre appears to be perceived independently of its pitch context.

To test the differences between effects for the corresponding pitch-relevant and timbre-relevant conditions, we performed a number of additional t tests. The first test examined the magnitude of interference produced by relevant-attribute context changes. This effect was significantly larger for pitch-relevant trials than for timbre-relevant trials, $t(17) = 2.424$, p

$< .05$. Thus, pitch and timbre are different in the extent to which they were coded in relation to the relevant-attribute values of context tones. In contrast, the magnitude of the effect of different-attribute context changes did not differ between pitch-relevant and timbre-relevant trials, $t(17) = -0.813$, $p > .05$. Neither pitch nor timbre recognition was subject to interference from irrelevant-attribute context changes.

To summarize, two main findings emerged from this experiment. The first is the different-attribute independence of timbre and pitch. Changing the pitches of tones surrounding the target tone did not interfere with memory for the timbre of the target, and changing the timbres of tones surrounding the target tone did not interfere with memory for the pitch of the target tone. (This independence is represented by the absence of diagonal lines in Figure 1.) Thus, it appears that pitch memory is unaffected by changing the timbres of contextual tones, which is consistent with the findings of Semal and Demany (1991). The present study also found the complementary kind of independence for timbre. These results suggest that the perceptual encoding of pitch and timbre has an integrity that is unaffected by the other attribute of the tones in the context. In this connection, note that the different-attribute independence found may depend on the rate of tone presentation. The studies of J. Smith et al. (1982) and Singh (1987) showed that timbre and pitch each contribute to stream segregation. If changing the pitch or timbre context results in a different organization of the standard and comparison sequences in the present kind of memory paradigm, then this might interfere with both pitch and timbre memory performance. The presentation rate of the present study, however, is considerably slower (intertone interval of 360 ms) than that typically used in streaming studies; the intertone interval in the study of J. Smith et al. (1982) was 250 ms, and the intertone interval in the study of Singh (1987) was 150 ms.

The second finding concerns same-attribute independence. As would be expected given the pervasive tendency to encode pitch in relation to other pitches, we found that changing the pitches of context tones interfered with memory for the pitch of the target tone. (This lack of independence is represented by horizontal lines between pitch values in Figure 1.) This was true even though the target tone always appeared in a predictable temporal position. The complementary result, however, was not found for timbre. Changing the timbres of the context tones did not interfere with memory for the timbre of the target tone. (This independence is represented by the absence of horizontal lines between timbre values in Figure 1.) It seems, then, that timbres may be encoded more in absolute terms (perhaps by identifying the instruments or some distinctive characteristics of the timbres; see Krumhansl, 1989) independent of their perceptual differences from the timbres of neighboring tones. How can this result be reconciled with the ability of Ehresman and Wessel's (1978) subjects to complete timbre analogies which required the abstraction of timbre intervals? A crucial difference between the studies may be that whereas their tones all had the same pitch, the tones in the present experiment varied in pitch. Thus, the pitch variation may prevent the perception of relations between successive timbres, which would seem prerequisite to perceiving timbre patterns and their transformations. This possibility was tested in the next experiment.

Experiment 3

This experiment further investigated whether an attribute of a tone is encoded in relation to that attribute of contextual tones. The previous experiment found that pitch was encoded relatively, but timbre was not. All of the sequences of that experiment, however, contained varying pitches and varying timbres. Apparently, the variations in timbre did not prevent pitch from being coded relatively, but variations in pitch may have prevented timbre from being coded relatively. This emphasis on the relative coding of these attributes reflects its importance for melodic patterns in music and whether analogous timbre patterns are possible. To test the effect of irrelevant attribute variations, in this final experiment we compared trials in which the irrelevant attribute was varied with trials in which the irrelevant attribute was held constant. For example, on some pitch-relevant trials, the sequences contained varying timbres, and on others the same timbre was used throughout. Analogously, on some timbre-relevant trials, the sequences contained varying pitches, and on others the same pitch was used throughout. The primary issue of interest was whether, when pitch was constant, timbres were perceived in relation to surrounding timbres, as reflected in interference of changing the timbres of context tones. The design of this experiment also allowed a second issue to be addressed. That issue was whether the complexity of the irrelevant attribute affects memory for the relevant attribute. In other words, is timbre memory better when pitch is constant than when it is varied, and is pitch memory better when timbre is constant than when it is varied? We refer to a difference of this sort as a *different-attribute complexity* effect.

Again, subjects listened to pairs of seven-tone sequences in which the fourth tone was the designated target. They were required to compare the target tones in the standard and comparison sequences in terms of the relevant attribute (pitch or timbre). Figure 5 illustrates the pitch-relevant conditions; the timbre-relevant conditions were constructed analogously. All of the examples shown are different trials as indicated by the change of the pitch of the target tone in the fourth serial position. On the first kind of trial, corresponding context tones in the standard and comparison sequences had the same pitches (no relevant-attribute context change), and the same timbre was used throughout (irrelevant-attribute constant). On the second kind of trial, corresponding context tones in the standard and comparison sequences did not have the same pitches (relevant-attribute context change), and the same timbre was again used throughout (irrelevant-attribute constant). The third kind of trial was like the first (no relevant-attribute context change), except the timbres varied (irrelevant-attribute varied). The fourth kind of trial was like the second (relevant-attribute context change), except the timbres varied (irrelevant-attribute varied). The third and fourth trial types are identical to two conditions of Experiment 2: no

Relevant Attribute: Pitch

Relevant-attribute	Irrelevant-attribute	Standard	Comparison
No Context Change	Constant	$P_1 P_8 P_4 \mathbf{P_3} P_2 P_7 P_6$ $T_1 T_1 T_1 \mathbf{T_1} T_1 T_1 T_1$	$P_1 P_8 P_4 \mathbf{P_5} P_2 P_7 P_6$ $T_1 T_1 T_1 \mathbf{T_1} T_1 T_1 T_1$
Context Change	Constant	$P_1 P_8 P_4 \mathbf{P_3} P_2 P_7 P_6$ $T_1 T_1 T_1 \mathbf{T_1} T_1 T_1 T_1$	$P_8 P_6 P_1 \mathbf{P_5} P_4 P_7 P_2$ $T_1 T_1 T_1 \mathbf{T_1} T_1 T_1 T_1$
No Context Change	Varied	$P_1 P_8 P_4 \mathbf{P_3} P_2 P_7 P_6$ $T_3 T_4 T_6 \mathbf{T_2} T_8 T_5 T_1$	$P_1 P_8 P_4 \mathbf{P_5} P_2 P_7 P_6$ $T_3 T_4 T_6 \mathbf{T_2} T_8 T_5 T_1$
Context Change	Varied	$P_1 P_8 P_4 \mathbf{P_3} P_2 P_7 P_6$ $T_3 T_4 T_6 \mathbf{T_2} T_8 T_5 T_1$	$P_8 P_6 P_1 \mathbf{P_5} P_4 P_7 P_2$ $T_3 T_4 T_6 \mathbf{T_2} T_8 T_5 T_1$

Figure 5. Example sequences illustrating the pitch-relevant conditions of Experiment 3; timbre-relevant conditions were constructed analogously. (The target zone was in the fourth serial position. The four conditions were produced by manipulating two variables independently: whether the pitches in the context were changed and whether the timbre was constant throughout the sequences or was varied. All of the sample trials illustrated are different trials.)

context change and relevant-attribute context change, respectively. The experiment examines whether the interference of a relevant-attribute context change depended on whether the irrelevant attribute was constant or varied and whether memory was affected by the complexity of the irrelevant attribute.

Method

Subjects. Subjects were 16 members of the Cornell University community who were selected according to the same criteria as in Experiment 1. They were paid for participating in the 1-hr experiment.

Apparatus. The apparatus was the same as in Experiment 1.

Stimulus materials. The tones used in the stimulus sequences had the same pitch and timbre values as in Experiment 2. Again, each trial consisted of a pair of seven-tone sequences in which the fourth (target) tone was either the same or different on the relevant attribute (pitch or timbre). Table 3 shows the conditions of the experiment. There were four conditions in which pitch was the relevant attribute and four conditions in which timbre was the relevant attribute. These four conditions were formed by crossing 2 two-valued factors. The first factor was whether the relevant-attribute context was the same in the standard and comparison sequences. The second factor was whether the irrelevant-attribute was constant or varied. When timbre was held constant, it was always the piano timbre; when it was varied, the values were randomly selected without replacement from the set of timbres used in Experiment 2. When pitch was held constant, it was always C5 (523 Hz); when it was varied, the values were randomly selected without replacement from the set of pitches used in Experiment 2.

Procedure. The instructions were the same as in Experiment 2. After reading the instructions, each subject was presented a training block of 16 practice trials with feedback (2 trials for each of the four conditions for each relevant attribute). The experimenter was present for this block to answer questions and to see if each subject understood the instructions correctly. After this was completed, the experimenter left the room, and each subject completed the four blocks of experimental trials. Subjects were free to take short breaks between blocks.

Each subject was presented two blocks of trials for each of the two relevant attributes (pitch and timbre) in an ordering counterbalanced between subjects. Each block started with 4 practice trials with feedback, continued with 4 practice trials without feedback, and concluded with 64 experimental trials. The experimental trials were a random ordering of 16 each (8 same and 8 different) for the four conditions for the relevant attribute.

Results and Discussion

Table 3 presents the results for both the pitch-relevant and timbre-relevant conditions. The MOC curves were based on ratings of 16 same and 16 different trials. For the pitch-relevant conditions, there was a significant main effect of relevant-attribute context, $t(15) = 4.512$, $p < .001$. Average performance was significantly better when the pitch context was not changed between the standard and comparison sequences (.814) than when it was the same (.710). The magnitude of the effect did not depend on whether the timbre varied or was constant, $t(15) = 1.069$, $p > .05$. In other words, the degree to which pitch was perceived relatively was the same for constant- and varied-timbre trials. Nor was there a main effect of whether the irrelevant attribute was constant or varied, $t(15) = 0.420$, $p > .05$. In particular, for trials on which the pitch context was unchanged, performance on varying-timbre trials (.810) was not significantly different from performance on constant-timbre trials (.826), $t(15) = 1.122$, $p > .05$. Thus, recognizing the target pitch was unaffected by whether the timbres were constant or varied.

For the timbre-relevant conditions, the effect of relevant-attribute context changes differed between pitch-constant and pitch-varying trials. There was a significant interaction between relevant-attribute context and whether the irrelevant-attribute context was constant or varied, $t(15) = 3.593$, $p < .01$. Indeed, only when a constant pitch was used throughout the standard and comparison sequences was there a significant difference between the no-relevant-attribute context-change

Table 3
Experiment 3: Conditions and Results

Relevant attribute	Irrelevant attribute	Area under memory-operating characteristic curve
	Relevant attribute: Pitch	
No context change	Constant	.826
Context change	Constant	.705
No context change	Varied	.801
Context change	Varied	.714
	Relevant attribute: Timbre	
No context change	Constant	.873
Context change	Constant	.809
No context change	Varied	.839
Context change	Varied	.865

condition (.873) and the relevant-attribute context-change condition (.809), $t(15) = 3.409$, $p < .01$. When pitch was varied, the effect of changing the timbre context was not significant (and even tended in the opposite direction, .839 vs. .865), $t(15) = -1.748$, $p > .05$. Thus, timbre appears to be perceived relatively only when pitch was constant. Overall, the relevant-attribute context did not have a significant main effect, $t(15) = 1.631$, $p > .05$. Nor was the main effect of whether pitch was constant or varied significant, $t(15) = -0.677$, $p > .05$. For trials on which the timbre context was unchanged, however, performance on constant pitch trials (.873) was marginally better than performance on varied-pitch trials (.839), $t(15) = 2.052$, $p = .058$. Thus, recognizing the target timbre in an unchanged timbre context was somewhat harder for pitch-varying sequences than for pitch-constant sequences.

We performed additional analyses to compare the magnitude of the effects across the corresponding pitch-relevant and timbre-relevant trials. The first analysis examined the magnitude of the effect of same-attribute context changes. This effect was significantly larger for pitch-relevant trials than for timbre-relevant trials, $t(15) = 3.002$, $p < .01$. Moreover, this effect depended on whether the irrelevant attribute was constant or varied. The difference in the effect between pitch-relevant and timbre-relevant conditions was significant only when the irrelevant attribute was varied, $t(15) = 3.167$, $p < .01$, but not when the irrelevant attribute was constant, $t(15) = 1.538$, $p > .05$. The second analysis examined the magnitude of the effect of whether the irrelevant attribute was constant or varied. The magnitude of the effect did not differ between pitch-relevant and timbre-relevant trials, $t(15) = 0.971$, $p > .05$. This was true also in separate analyses of no-relevant-attribute context-change conditions, $t(15) = -0.402$, $p < .05$, and in relevant-attribute context-change conditions, $t(15) = 1.350$, $p > .05$.

Of central interest in this experiment was the finding that timbre was perceived in relation to other timbres when pitch was held constant. When the same pitch was used throughout the sequences, changing the timbres of the context tones interfered with memory for the timbre of the target tone. (This lack of independence is represented by horizontal lines between timbre values in Figure 1.) This kind of interference was not found, however, in either this or the previous experiment when pitch was varied. These pitch variations apparently interfered with the relative perception of timbre. This result is consistent with Miller and Carterette's (1975) finding that pitch variations predominated over timbre variations in similarity judgments, suppressing the contribution of timbre relations. It is also consistent with the suggestion made earlier that the ability of subjects in Ehresman and Wessel's (1978) study to complete timbre analogies may have depended on the fact that all of the tones in their study had the same pitch. Thus, the relative perception of timbre appears to be unstable and subject to interference from pitch variations. In contrast, this and the previous experiment showed that the relative perception of pitch is a robust phenomenon. (This lack of independence is represented by horizontal lines between pitch values in Figure 1.) The present experiment showed that the interference of changing the pitch of context tones on pitch memory was equally strong whether the timbre was varied or constant. In terms of how successive tones are perceived, then, pitch and timbre show quite different degrees of relative perception. A related finding concerns the complexity effect of the irrelevant attribute, and again the results for pitch and timbre were different. Memory for the target pitch was no different when the timbre varied from when it was constant. In contrast, memory for the target timbre was marginally worse when pitch varied than when it was constant. Thus, these results show another asymmetry between pitch and timbre; pitch perception was less affected by timbre variations than timbre perception was by pitch variations.

General Discussion

Figure 6 schematically summarizes the results of the three experiments. In the first experiment we used the Garner (1974) classification tasks to find that the timbre and pitch of a single tone interact. This is indicated in Figure 6 by the vertical lines connecting the corresponding timbre and pitch values. In the experiment, the interaction made uncorrelated variation on the irrelevant attribute difficult to filter out and correlated variation on the irrelevant attribute beneficial. This finding is consistent with previous studies by Beal (1985), Crowder (1989), and Melara and Marks (1990a, 1990b, 1990c), all of which showed perceptual interactions between pitch and timbre. Although these two attributes are defined and can be manipulated independently, they are not perceived independently. Thus, musical pitch and timbre are like many other perceptual attributes, such as hue and brightness of color, that interact in perception.

The perceptual interaction between timbre and pitch of single tones does not imply, however, that it is impossible to abstract and compare pitches of tones with different timbres or timbres of tones with different pitches. The low error rates in this experiment demonstrated this ability, but the increased reaction times in tasks requiring this information to be abstracted indicates that this process requires additional time. The differences found by Beal (1985) between musicians and nonmusicians suggest that musical experience is also important. Comparing pitches of different timbres probably depends on determining the fundamental frequency of the harmonically rich tones. The fundamental frequency is the greatest common divisor of all of the frequency components (harmonics) and corresponds to the pitch that is heard. It is more

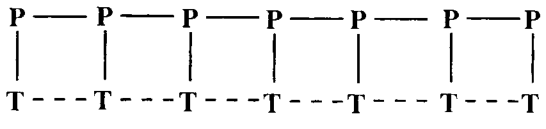

Figure 6. Results of the three experiments. (The vertical lines indicate the interaction found in Experiment 1 between the pitch and timbre of a single tone. The horizontal lines between successive pitches indicate the strong tendency found in Experiments 2 and 3 to perceive pitch in relation to other pitches in the context. The dashed horizontal lines between successive pitches indicate the weak tendency found in Experiment 3 [only when pitch is constant] to perceive timbre in relation to other timbres in the context.)

difficult to speculate on the process through which timbres of different pitches are compared. The acoustic properties of tones produced by a single instrument depend on a number of factors, including the pitch, loudness, and the manner of playing. For example, the harmonic structure of a violin tone depends on whether the tone is in the low or high pitch register, whether it is soft or loud, and whether the bow is near or far from the bridge. Despite the acoustic differences, listeners can identify the instrument, suggesting that some kind of underlying acoustic invariants contribute to the perception of timbre. At present, this difficult problem has not received systematic analysis from a perceptual point of view.

In Figure 6, the horizontal lines connecting the different pitch values indicate that pitch is perceived in relation to other pitches in the context. For simplicity, only the connections between successive pitches are shown, but relations between more distant pitches are probably also perceived. In the present study, the tendency to perceive pitch in relation to other pitches was manifested in interference from changing the pitch of context tones (tones preceding and following the target tone). Memory for the target pitch was consistently better when the pitches of the context tones were unchanged (a lack of same-attribute independence; Experiments 2 and 3). This was true even though the target tone always appeared in a fixed temporal position in the relatively short sequences used. The magnitude of the interference effect was unaffected by whether the timbre was constant or varied throughout the sequences (no interaction with different-attribute complexity; Experiment 3). That the perception of pitch is largely unaffected by the timbre of the context was further substantiated by the equal performance on pitch recognition in timbre-constant and timbre-varying sequences when the pitch context was unchanged (no effect of different-attribute complexity; Experiment 3). Thus, the relative perception of pitch is robust under a variety of conditions involving timbre manipulations. Relative pitch perception is extremely important in music perception. It subserves the perception of melodic and harmonic patterns, the abstraction of tonality, and the recognition of melodies under a variety of transformations such as transposition and inversion (for summaries, see Dowling & Harwood, 1986; Krumhansl, 1990).

The present experiments showed that the relative perception of timbre is considerably weaker, as indicated by the dashed lines between successive timbres in Figure 6. Only when pitch was held constant did changing the timbres of context tones interfere with memory for the target timbre (a lack of same-attribute independence; Experiment 3). When pitch varied, changing the timbres of the context tones had no effect (same-attribute independence; Experiment 2 and 3), suggesting that the pitch variations interfered with the perception of timbre relations. A related result was that timbre memory was impaired by variations in pitch; memory for the target timbre, when the timbre context was not changed, was somewhat worse in pitch-varying than in pitch-constant sequences (different-attribute complexity; Experiment 3). The only result showing that timbre was unaffected by pitch was that memory for the target timbre was unaffected by changing the pitch context (different-attribute independence; Experiment 2).

Implicit in this discussion of whether timbre is perceived in absolute or relative terms is a bias that stems from considering timbre from a musical point of view. Unless the relations between timbres are perceived, it seems unlikely that attempts to establish hierarchically organized patterns of timbres (Lerdahl, 1987) or to effect transformations of timbre patterns such as transposition and inversion (Slawson, 1985) will be successful. Together, the present results on timbre suggest that it may be difficult to perceive patterns of timbre variations in music unless pitch variations are highly controlled. Thus, the musical function of timbre may be subject to perceptual limitations. In this connection, note that Slawson (1985) has been concerned with timbres with vowellike qualities, which may be subject to different perceptual principles from the musical instrument timbres used in the present study. The magnitude of the differences between timbres may also be an important factor. The timbres used in these experiments were distinctive and chosen to span the space of musical instrument timbres. It may be easier to perceive relations between timbres with more subtle differences, such as those that can be synthesized electronically. In addition, it would be interesting to study whether timbre relations would be perceived more easily if the pitch variations constituted a coherent melody as opposed to the random pitch sequences used in these experiments. The question of how musical timbre can be controlled, manipulated, and perceived invites further analysis from the perspectives of acoustics, music theory and composition, and experimental psychology.

References

Beal, A. L. (1985). The skill of recognizing musical structures. *Memory & Cognition, 13,* 405–412.

Benade, A. H. (1976). *Fundamentals of musical acoustics.* New York: Oxford University Press.

Berger, K. W. (1963). Some factors in the recognition of timbre. *Journal of the Acoustical Society of America, 36,* 1888–1891.

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

Bregman, A. S., & Campbell, J. (1971). Primary auditory stream segregation and perception of order in rapid sequences of tones. *Journal of Experimental Psychology, 89,* 244–249.

Cogan, R. (1984). *New images of musical sound.* Cambridge, MA: Harvard University Press.

Crowder, R. G. (1989). Imagery for musical timbre. *Journal of Experimental Psychology: Human Perception and Performance, 15,* 472–478.

Deutsch, D. (1972). Mapping of interactions in the pitch memory store. *Science, 175,* 1020–1022.

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

Dowling, W. J., & Harwood, D. L. (1986). *Music cognition.* Orlando, FL: Academic Press.

Ehresman, D., & Wessel, D. (1978). Perception of timbral analogies. *Rapports IRCAM, 13/78.* Paris: Centre Georges-Pompidou.

Erickson, R. (1976). *Sound structure in music.* Berkeley: University of California Press.

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

Garner, W. R. (1981). The analysis of unanalyzed perceptions. In M.

Kubovy & J. R. Pomerantz (Eds.), *Perceptual organization* (pp. 119–139). Hillsdale, NJ: Erlbaum.

Grey, J. M. (1977). Multidimensional perceptual scaling of musical timbres. *Journal of the Acoustical Society of America, 61,* 1270–1277.

Grey, J. M. (1978). Timbre discrimination in musical patterns. *Journal of the Acoustical Society of America, 64,* 467–472.

Grey, J. M., & Gordon, J. W. (1978). Perceptual effects of spectral modifications of musical timbres. *Journal of the Acoustical Society of America, 63,* 1493–1500.

Handel, S. (1989). *Listening: An introduction to the perception of auditory events.* Cambridge, MA: MIT Press.

Iverson, P., & Krumhansl, C. L. (1991). Measuring similarity of musical timbres. *Journal of the Acoustical Society of America, 89*(Suppl. 2), 1988.

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

Jones, M. R., Kidd, G., & Wetzel, R. (1981). Evidence for rhythmic attention. *Journal of Experimental Psychology: Human Perception and Performance, 7,* 1050–1072.

Jones, M. R., Summerell, L., & Marshburn, E. (1987). Recognizing melodies: A dynamic interpretation. *Quarterly Journal of Experimental Psychology, 39A,* 89–121.

Kendall, R. A. (1986). The role of acoustic signal partitions in listener categorization of musical phrases. *Music Perception, 4,* 185–214.

Krumhansl, C. L. (1979). The psychological representation of musical pitch in a tonal context. *Cognitive Psychology, 11,* 346–374.

Krumhansl, C. L. (1989). Why is musical timbre so hard to understand? In S. Nielzen & S. Olsson (Eds.), *Structure and perception of electroacoustic sound and music* (pp. 43–53). Amsterdam: Elsevier.

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., & Schmuckler, M. A. (1986). The *Petroushka* chord: A perceptual investigation. *Music Perception, 4,* 153–184.

Lerdahl, F. (1987). Timbral hierarchies. *Contemporary Music Review, 2,* 135–160.

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 and Performance, 16,* 398–414.

Miller, J. R., & Carterette, E. C. (1975). Perceptual space for musical structures. *Journal of the Acoustical Society of America, 58,* 711–720.

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

Monahan, C. B., Kendall, R. A., & Carterette, E. C. (1987). The effect of melodic and temporal contour on recognition memory for pitch change. *Perception & Psychophysics, 41,* 576–600.

Nakamura, T. (1987). The communication of dynamics between musicians and listeners through musical performance. *Perception & Psychophysics, 41,* 525–553.

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

Pomerantz, J. R., & Pristach, E. A. (1989). Emergent features, attention, and perceptual glue in visual form perception. *Journal of Experimental Psychology: Human Perception and Performance, 15,* 635–649.

Risset, J-C., & Wessel, D. L. (1982). Exploration of timbre by analysis and synthesis. In D. Deutsch (Ed.), *The psychology of music* (pp. 26–58). New York: Academic Press.

Saldanha, E. L., & Corso, J. F. (1964). Timbre cues and the identification of musical instruments. *Journal of the Acoustical Society of America, 36,* 2021–2026.

Semal, C., & Demany, L. (1991). Dissociation of pitch from timbre in auditory short-term memory. *Journal of the Acoustical Society of America, 89,* 2404–2410.

Singh, P. G. (1987). Perceptual organization of complex-tone sequences: A tradeoff between pitch and timbre? *Journal of the Acoustical Society of America, 82,* 886–899.

Slawson, W. (1985). *Sound color.* Berkeley, CA: University of California Press.

Smith, J., Hausfeld, S., Power, R. P., & Gorta, A. (1982). Ambiguous musical figures and auditory streaming. *Perception & Psychophysics, 32,* 454–464.

Smith, K. C., & Cuddy, L. L. (1989). Effects of metric and harmonic rhythm on the detection of pitch alterations in melodic sequences. *Journal of Experimental Psychology: Human Perception and Performance, 15,* 457–471.

Swets, J. A. (1973). The relative operating characteristic in psychology. *Science, 182,* 990–1000.

Von Bismarck, G. (1974). Timbre of steady sounds: A factorial investigation of its verbal attributes. *Acoustica, 30,* 146–159.

Wedin, L., & Goude, G. (1972). Dimension analysis of the perception of instrumental timbre. *Scandinavian Journal of Psychology, 13,* 228–240.

Wessel, D. L. (1973). Psychoacoustics and music. *Bulletin of the Computer Arts Society, 30,* 1–2.

Received March 15, 1991
Revision received July 31, 1991
Accepted August 2, 1991 ■