Journal of Experimental Psychology:
Human Perception and Performance
2001, Vol. 27, No. 1, 128–140

0096-1523/01/$5.00 DOI: 10.1037//0096-1523.27.1.128

Illusory Conjunctions of Pitch and Duration in Unfamiliar Tone Sequences

William Forde Thompson
York University

Michael D. Hall
University of Nevada, Las Vegas

Jeff Pressing
University of Melbourne

In 3 experiments, the authors examined short-term memory for pitch and duration in unfamiliar tone sequences. Participants were presented a target sequence consisting of 2 tones (Experiment 1) or 7 tones (Experiments 2 and 3) and then a probe tone. Participants indicated whether the probe tone matched 1 of the target tones in both pitch and duration. Error rates were relatively low if the probe tone matched 1 of the target tones or if it differed from target tones in pitch, duration, or both. Error rates were remarkably high, however, if the probe tone combined the pitch of 1 target tone with the duration of a different target tone. The results suggest that illusory conjunctions of these dimensions frequently occur. A mathematical model is presented that accounts for the relative contribution of pitch errors, duration errors, and illusory conjunctions of pitch and duration.

Pitch and duration are critical dimensions in the music of many cultures throughout the world. In Western tonal music, compositions involve a relatively small set of discrete pitch and duration categories, and listeners are highly sensitive to patterns of pitch and duration (i.e., melody and rhythm). The distinct categories used to classify different pitches and durations in compositions and our sensitivity to the properties associated with patterns of pitch and duration suggest that there are important mechanisms for encoding these properties.

The initial processing of music may involve two basic stages. In the first stage, acoustic dimensions such as pitch, duration, timbre, location, and loudness may be registered by functionally independent channels (Garner, 1974; Livingstone & Hubel, 1987). In the second stage, these acoustic dimensions may be recombined to yield the experience of a unified auditory event. If a processing error occurs at the latter stage, listeners may experience an *illusory conjunction* of features from different acoustic events (Deutsch, 1986; Thompson, 1994).

William Forde Thompson, Department of Psychology, Atkinson Faculty of Liberal and Professional Studies, York University, Toronto, Ontario, Canada; Michael D. Hall, Department of Psychology, University of Nevada, Las Vegas; Jeff Pressing, Department of Psychology, University of Melbourne, Melbourne, Victoria, Australia.

This research was supported by Grant 46675-97 from the Natural Sciences and Engineering Research Council of Canada and by an institutional training grant in speech and hearing sciences awarded to the University of Washington (NIDCD T32 DC00033-07).

Doug Gifford, Randy Jamieson, Rani Neeru, and Murray Stainton provided technical assistance. We thank Greg Ashby, Carol Fowler, Bill Prinzmetal, and an anonymous reviewer for comments on earlier versions of this article.

Correspondence concerning this article should be addressed to William Forde Thompson, Department of Psychology, Atkinson Faculty of Liberal and Professional Studies, York University, 4700 Keele Street, Toronto, Ontario M3J 1P3, Canada. Electronic mail may be sent to billt@yorku.ca.

The analysis of music into distinct structural dimensions is an approach supported by theory, pedagogy, and compositional practice. For example, treatises on music psychology (e.g., Deutsch, 1999), ear training (e.g., Hindemith, 1946), and harmony (e.g., Forte, 1974) all assume an operational independence between pitch, duration, and other dimensions of music. Moreover, traditional pedagogical exercises address pitch and rhythmic skills separately, implying separate mechanisms for processing these dimensions. Finally, in Western tonal composition, duration and pitch often vary independently of each other. For example, in the *Art of the Fugue,* J. S. Bach uses mensuration canons, in which the same melody is presented at different speeds simultaneously. Similarly, in the widely used Schenkerian reduction technique for analyzing tonal works (e.g., Forte & Gilbert, 1982), pitch structure is considered to operate independently of surface temporal relations. Thus, pitch and duration often operate independently in music. How sensitive are listeners to the manner in which they are combined?

In this investigation, we examined listeners' sensitivity to combinations of pitch and duration by assessing the possibility that illusory conjunctions of pitch and duration occur in short-term memory. Listeners were asked to remember combinations of pitch and duration in short unfamiliar tone sequences. Memory errors were then examined for evidence of illusory conjunctions. We also conducted mathematical modeling to estimate the precise rate at which pitch and duration were incorrectly integrated.

Encoding Auditory Dimensions

The question of how auditory dimensions are encoded (e.g., pitch, duration, timbre, location, loudness) may be addressed by assessing whether they interact in perception or memory or whether they are registered by functionally independent channels (Garner, 1974; Livingstone & Hubel, 1987). On the basis of results of speeded classification tasks, Melara and Marks (1990a, 1990b, 1990c) proposed an interactive multichannel model of auditory processing. They conceived of auditory processing as a bank of channels, each targeted to a specific primary perceptual dimension (e.g., frequency, loudness). Judgments of one dimension were sometimes affected by variation in irrelevant dimensions, illustrating leakage or cross-talk between channels (see also Krumhansl & Iverson, 1992).

According to some researchers, the combined effects of pitch and rhythm suggest that these dimensions are encoded integrally. Jones, Boltz, and Kidd (1982) found that recognition for the pitch of a tone in a melodic context is better if the tone was rhythmically accented than if it was not rhythmically accented. Moreover, memory for melodic sequences is poor if the pitch pattern and rhythm imply incompatible metric groupings (Boltz & Jones, 1986; Deutsch, 1980). These and other findings have led to the proposal that pitch pattern and rhythm are encoded as a single unit in perception and memory (see also Jones, 1987; Jones, Summerell, & Marshburn, 1987).

Other evidence, however, suggests that pitch pattern and rhythm are processed separately. First, pitch pattern and rhythm make statistically independent contributions to judgments of melodic similarity (Monohan & Carterette, 1985) and phrase completion (Palmer & Krumhansl, 1987a, 1987b). Second, listeners have difficulty remembering how pitch patterns and rhythms are combined in short melodies, even when memory for the individual components is good (Thompson, 1994, Experiment 4). Third, double dissociations of pitch pattern and rhythm have been documented in brain-damaged patients, suggesting a neural dissociation between these dimensions (Peretz, 1993; Peretz & Kolinsky, 1993; Peretz & Morais, 1989).

Finally, in pilot work conducted for this study, listeners were presented unfamiliar seven-tone sequences followed by a probe tone. In a pitch-search condition, listeners judged whether the probe tone matched any of the target tones in pitch. In a duration-search condition, listeners judged whether the probe tone matched any of the target tones in duration. In half of the trials within each condition, the irrelevant feature was varied across target tones; in the other half, the irrelevant feature was held constant. Memory for one feature was not influenced by variation in the irrelevant feature, suggesting that pitch and duration are processed separately in short-term memory.

Illusory Conjunctions

If pitch and duration are processed independently in short-term memory for novel tone sequences, it may be difficult to remember how these features are combined. In such circumstances, listeners may be susceptible to illusory conjunction errors; that is, they may falsely remember the pitch of tone combined with the duration of a different tone. It is well established that illusory conjunctions of visual features occur, and a similar phenomenon may occur for auditory features. Such errors suggest that once features are registered, a subsequent process is required to integrate features and perceive objects (e.g., Ivry & Prinzmetal, 1991; Prinzmetal, Henderson, & Ivry, 1995; Treisman, 1991; Treisman & Schmidt, 1982).

A number of tasks have been used to demonstrate the occurrence of illusory conjunctions in vision. One task involves asking participants to detect the presence or absence of a conjunction of two features in a briefly presented visual array. Treisman and Schmidt (1982) used this task to reveal illusory conjunctions of color and shape (e.g., perceiving a red *O* or a green *X* when presented a red *X* and a green *O*) and illusory conjunctions of size and solidity (filled vs. outlined shapes). As would be expected from a system with limited processing capacity, illusory conjunction rate was found to increase when attention is overloaded.

Evidence of illusory conjunctions in vision has been critical to the development of a unified theory of visual object perception (e.g., Treisman, 1991; Treisman & Gelade, 1980; Treisman & Gormican, 1988; Treisman & Schmidt, 1982; Treisman & Souther, 1986). According to feature-integration theory (FIT), individual features are first preattentively registered in parallel from objects in a visual array and tagged with respect to the location where they occur. Integration of these features then is argued to occur through a serial process of focusing attention at specific locations.

A number of researchers have questioned some of the assumptions of FIT, including that of parallel and serial processing of items (e.g., Duncan & Humphreys, 1989; Wolfe, Cave, & Franzel, 1989) and the role of attention in the integration process (Ashby, Prinzmetal, Ivry, & Maddox, 1996; Prinzmetal et al., 1995). As an alternative to FIT, Ashby et al. (1996) proposed that illusory conjunctions arise from uncertainty about the location of individual features. Support for location-uncertainty theory was based on findings that illusory conjunction errors increased as the distance between display items (colored letters) decreased. Other models of illusory conjunction phenomena have been proposed by Wolford (1975) and by Maddox, Prinzmetal, Ivry, and Ashby (1994). Although such models embody different assumptions about object perception, all share the assumption that illusory conjunctions occur and that they result from errors in a process of feature integration.

Evidence for illusory conjunctions of auditory features is limited. Efron and Yund (1974) and Deutsch (1975) reported illusory conjunctions between the pitch of a tone presented at one ear and the intensity or location of a tone presented at the other ear. Cutting (1976) found that listeners may mistakenly combine the features of two dichotically presented phonemes and report illusory phonemes. Finally, in a study by Hall, Pastore, Acker, and Huang (2000), listeners searched arrays of spatially distributed, simultaneously presented tones for the presence or absence of a cued musical pitch, a cued instrument timbre, or a cued conjunction of both properties. These tasks were used because they are analogous to the visual-search tasks used by Treisman and Schmidt (1982). The results, supported by mathematical modeling, provided evidence for the occurrence of illusory conjunctions of pitch and timbre.

Memory for Pitch and Duration: The Current Investigation

Previous research has established that listeners are not highly sensitive to the manner in which pitch and duration are combined in atonal music (Krumhansl, 1991) or in unfamiliar tone sequences (Thompson, 1994). These findings, along with neuropsychological evidence described earlier, raise the possibility that pitch and duration are processed separately and that illusory conjunctions between these dimensions may occur.

Thompson (1994, Experiments 1–3) presented listeners with sequences in which two different tones were played consecutively and repeatedly, in a randomly varying order. Such sequences were defined as *melodic textures*. Listeners were asked to indicate when they detected a change in texture. If a new pitch or a new duration was introduced into the texture, listeners easily perceived the texture change, even while performing a distraction task. However, if the texture change was defined by the manner in which pitch and duration were combined in the two tones, only attentive listeners could detect the change. This auditory illustration of a conjunction-search task differs from the visual-search tasks introduced by Treisman and Gelade (1980) in that items are distributed across time rather than spatial location. Thus, Thompson's (1994) results raise the possibility that illusory conjunctions may occur across temporal location as well as across spatial location.

The current investigation further examined the occurrence of illusory conjunctions of pitch and duration in short-term memory. In three experiments, we presented a short target sequence and then a probe tone. Listeners indicated whether the probe tone matched one or more of the target tones in both pitch and duration. To anticipate the results, error rates (false match responses) were remarkably high if the probe tone combined the pitch of one target tone with the duration of a different target tone. In contrast, error rates were comparatively low if the probe tone differed from target tones in an individual feature (i.e., in pitch and/or duration). A mathematical model of response errors was developed to estimate the precise incidence of illusory conjunctions and feature errors.

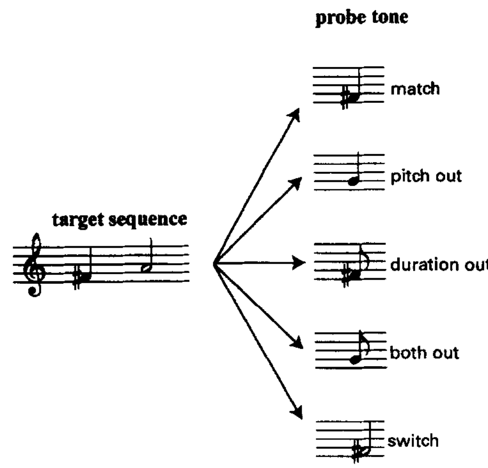

Figure 1. An illustration of a target sequence used in Experiment 1 and the probe tones used for each of the five conditions.

Experiment 1

We conducted Experiment 1 to assess short-term memory for combinations of pitch and duration and to evaluate the possibility that illusory conjunctions of these dimensions occur. Participants were presented two tones in sequence (called target tones) and then a probe tone. They were asked whether the probe tone matched one of the two target tones in both pitch and duration. That is, participants were required to search (in short-term memory) for a conjunction of pitch and duration. We presented two target tones in order to distinguish memory for individual dimensions (pitch and duration) from memory for how those dimensions are combined. In particular, the two target tones differed from each other in both pitch and duration; thus, illusory conjunctions were implicated when listeners mistakenly remembered the pitch of one target tone combined with the duration of the other target tone.

There were five conditions, which were defined by the relationship between the probe tone and the target tones. The probe tone either matched one of the target tones (match condition) or differed from both target tones in pitch (pitch-out condition), duration (duration-out condition), both pitch and duration (both-out condition), or the manner in which pitch and duration were combined (switch condition). In the latter condition, we constructed the probe tone by combining the pitch of one target tone with the duration of the other target tone. Thus, whereas both features of the probe tone were present in the target sequence, the probe tone in the switch condition did not match either target tone in both pitch and duration. Figure 1 provides the notation for one target sequence and the probe tones used for each of the five conditions.

We made three predictions regarding the errors made by participants. First, the proportion of errors should be lower if the probe tone differs from target tones in an individual feature (e.g., pitch-out, duration-out, or both-out conditions) than if the probe tone differs from target tones in how pitch and duration are combined (switch condition). Detection in the former cases requires only accurate registration of all target-tone features; detection in the latter case requires both accurate registration and accurate integration of target-tone features. Thus, the latter condition should result in an increased error rate because of the contribution of errors in feature integration (i.e., illusory conjunctions).

Second, the proportion of errors should be lower if the probe tone differs from all target tones in both pitch and duration (both-out condition) than if the probe tone differs from all target tones along just one of these dimensions (pitch-out and duration-out conditions). That is, confusions are more likely if target tones have one feature in common with the probe tone than if target tones have no features in common with the probe tone. Third, because listeners are expected to have at least some sensitivity to the manner in which pitch and duration are combined, the match condition should yield a higher (though similar) proportion of match responses than the switch condition. In both conditions, the pitch and duration of the probe tone are present in the target sequence. In the switch condition, however, these two features occur in different target tones.

Method

Participants. Participants were 20 adults from the York University community, ranging in age from 19 to 48 years ($M = 23.3$). No musical training was required to participate in the experiment. A survey revealed that participants had between 0 and 12 years of musical training ($M = 4.2$ years). All participants reported normal hearing.

Stimuli. Each trial consisted of four tones—two target tones surrounded by two anchor tones—followed by a pause of 5.0 s and then a probe tone. The anchor tones were 500 ms in duration and always had a pitch of middle C (C4). They were included because previous research suggests that illusory conjunctions for pairs of target items are increased if those items are flanked by irrelevant stimuli (Ashby et al., 1996; Treisman, 1982, Experiment 4). In all trials, the pitch of the anchor tones (C4) always differed from the pitch of both target tones and the pitch of the probe tone.

Anchor tones were separated from target tones by a pause of 500 ms. Target tones were separated from each other by a pause of either 0 or 1 s.

Tones were generated by a Roland Sound Canvas that was under the control of a Power Macintosh and was set to equal temperament tuning. The timbre of tones was fixed as the square-wave sound of the Roland Sound Canvas. The tempo of sequences was set to 120 beats per minute, or 500 ms per quarter note. Participants were allowed to adjust the intensity to a comfortable listening level between 60 and 70 dB SPL, and they heard the sequences through Sennheisser HD-480 headphones.

There were 16 sequences of target tones. The pitches of target tones were randomly varied between G3 and G4, with the constraint that neither target tone matched the anchor tones in pitch. Each target sequence was presented with each of five different probe tones (one for each of the five conditions described above), yielding a total of 80 trials. All sequences of target tones involved two different duration values: 500 ms and 250 ms, 250 ms and 125 ms, or 1,000 ms and 500 ms.

Procedure. After each presentation, participants indicated whether or not the probe tone matched one of the two target tones in both pitch and duration. Participants made their responses by using a mouse to select one of two response options that appeared on the computer screen: *YES* (the probe tone matched one of the target tones) or *NO* (the probe tone did not match any of the target tones). Practice trials, selected at random from the experimental trials, were provided to acquaint the participants with the task. Participants were allowed to quit the practice trials and begin the experimental trials at any time.

Results and Discussion

Figure 2 displays the proportion of *Yes* (i.e., match) responses for each of the five conditions. Match responses were subjected to a two-way analysis of variance (ANOVA), with repeated measures on the factors of separation of the target tones (0 and 1 s) and condition (match, pitch-out, duration-out, both-out, and switch). The ANOVA revealed no effect of separation and no interaction between separation and condition, but it did reveal a highly significant effect of condition, $F(4, 76) = 41.20$, $p < .001$. Figure 2 illustrates that the match and switch conditions were both associated with a relatively high incidence of match responses.

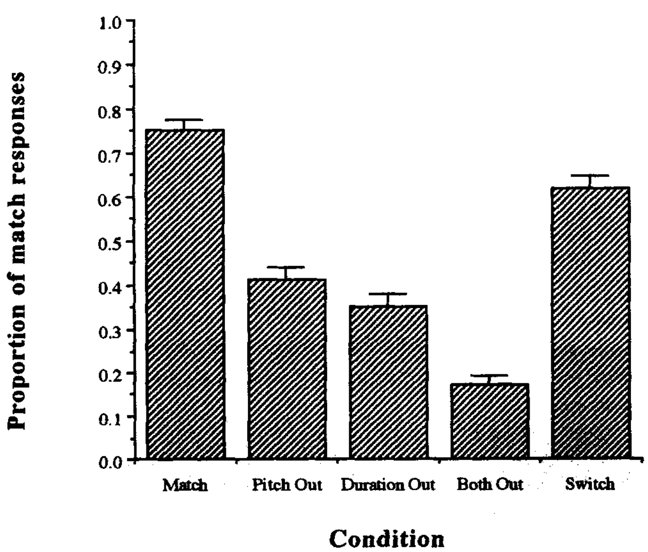

Figure 2. The proportion of errors for each of the five conditions in Experiment 1.

Planned comparisons confirmed the three predictions for Experiment 1. First, the proportion of false match responses was significantly lower for the duration-out and pitch-out conditions than for the switch condition, $F(1, 76) = 30.07$, $p < .001$. This finding confirms that listeners were significantly more sensitive to changes in individual features (pitch or duration) than to changes in how features were combined. More specifically, it suggests that memory for the target sequences often involved illusory conjunctions of pitch and duration.

Second, the proportion of false match responses was lower for the both-out condition than for the duration-out and pitch-out conditions, $F(1, 76) = 23.41$, $p < .001$. This finding indicates that participants were more sensitive to differences in two dimensions (pitch and duration) than in just one dimension.

Third, the proportion of false match responses for the switch condition was significantly lower than the proportion of (correct) match responses for the match condition, $F(1, 76) = 6.80$, $p < .05$. This finding suggests that listeners retained some sensitivity to the manner in which pitch and duration were combined. In both the switch and match conditions, the pitch and duration of the probe tone were present in target sequences; the two conditions are only differentiated by whether those features occurred in a single target tone (match condition) or in different target tones (switch condition).

Signal detection analyses further confirmed that listeners had difficulty encoding combinations of pitch and duration.[1] Sensitivity to differences between the probe tone and target tones was significantly poorer for the switch condition ($M\ d' = 0.42$) than for the pitch-out ($M\ d' = 1.03$), duration-out ($M\ d' = 1.19$), and both-out conditions ($M\ d' = 1.76$), $F(1, 57) = 46.76$, $p < .0001$. Thus, participants were significantly more sensitive to individual features in target tones than to how those features were combined. Nonetheless, the mean d' value in the switch condition was greater than zero, $t(19) = 3.61$, $p < .01$, indicating that listeners were not entirely insensitive to the manner in which pitch and duration were combined.

Experiment 2

The results of Experiment 1 indicated that listeners are remarkably insensitive to the manner in which pitch and duration are combined in two-tone target sequences. Moreover, the high proportion of match responses on the switch condition provides strong evidence for illusory conjunctions of pitch and duration. When the probe tone matched the pitch of one target tone and the duration of the other target tone (but did not match any one target tone), participants responded on 62% of trials that the probe tone matched an individual target tone in both pitch and duration. Experiment 2 was conducted to assess whether listeners also have difficulty remembering combinations of pitch and duration in

[1] Sensitivity to differences between the probe tone and target tones was evaluated by comparing the proportion of *no* responses on the pitch-out, duration-out, both-out, and switch conditions (hits) with the proportion of *no* responses on the match condition (false alarms). Proportions of 0 and 1 were converted to .05 and .95, respectively. Other strategies for dealing with proportions of 0 and 1 are discussed in Macmillan and Creelman (1991, chap. 1). Alternative strategies were examined, but none altered the test outcome.

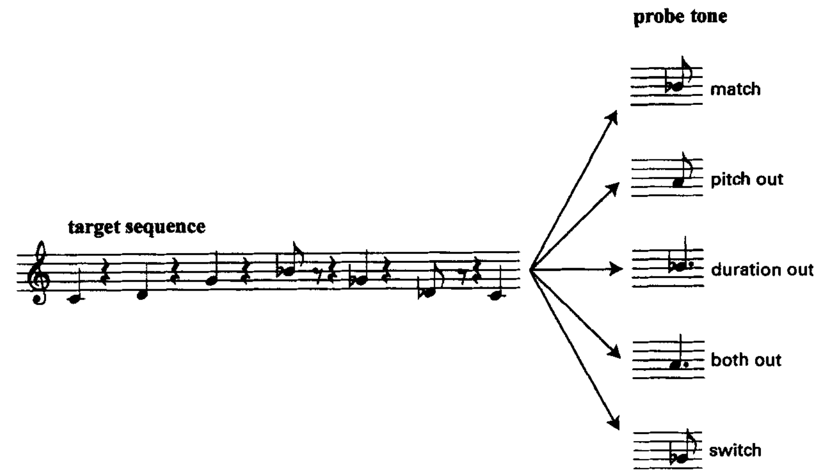

Figure 3. An illustration of a target sequence used in Experiment 2 and the probe tones used for each of the five conditions.

longer sequences. In this case, target sequences consisted of seven tones, which is a common phrase length in Western melodies (e.g., "Mary Had a Little Lamb," "Twinkle Twinkle Little Star"). Difficulty remembering combinations of pitch and duration in such sequences would suggest that memory for unfamiliar melodies may often involve illusory conjunctions of melodic features.

Method

Participants. Participants were 18 adults from the University of New South Wales student community, ranging in age from 18 to 23 years (M = 19.28). None of the participants had been tested in Experiment 1. A survey revealed that participants had between 0 and 10 years of musical training (M = 3.19 years). All participants reported normal hearing.

Stimuli. Each trial consisted of a seven-tone target sequence and then a probe tone. The interonset interval (IOI) between adjacent tones in the target sequence was 1,000 ms. The IOI between the last (seventh) target tone and the probe tone was 1,500 ms. Tones were generated in the same way as in Experiment 1. There were 16 target sequences. Each sequence started and ended on the same pitch. Except for the first and last target tones, all target tones differed in pitch and involved an upward melodic contour for the first four tones followed by a downward melodic contour for the last three tones. The starting pitch of target sequences varied randomly between G2 and D3, and the pitch range of sequences was between 7 and 11 semitones. Aside from these constraints, pitches were assigned at random.

Each target sequence consisted of a pattern of two durations (which is typical of Western melodic phrases): One of these duration values was assigned to five target tones and the other duration value was assigned to two target tones. Tone durations were one of the following: 125 ms (sixteenth note), 250 ms (eighth note), 500 ms (quarter note), or 750 ms (dotted quarter note). Aside from these constraints, durations were assigned at random. Note that the IOI between adjacent target tones was always greater than the duration of any target tone, leaving a pause between all adjacent target tones. Pauses were included to facilitate the encoding of individual tone durations.

Each of the 16 target sequences was presented with each of five different probe tones (i.e., 80 trials), thus defining the five conditions described in the *Method* section of Experiment 1 (match, pitch-out, duration-out, both-out, and switch). Figure 3 provides the notation for one target sequence and the probe tones used for each of the five conditions. For conditions in which the probe tone matched a target tone in pitch, duration, or both pitch and duration, the match never occurred on the first or last target tone (to avoid enhanced performance from primacy and recency effects). For conditions in which the probe tone matched a target tone in duration, the match in duration occurred with two target tones.[2]

Procedure. The procedure was identical to that used in Experiment 1.

Results and Discussion

Figure 4 displays the proportion of match responses for each of the five conditions. An ANOVA revealed a significant effect of condition, $F(4, 68) = 36.03$, $p < .001$, with the match and switch conditions associated with a comparatively high proportion of match responses.

Planned comparisons again confirmed the three predictions outlined in Experiment 1. First, the proportion of false match responses was significantly lower for the duration-out and pitch-out conditions than for the switch condition, $F(1, 68) = 30.89$, $p < .001$. This finding illustrates that participants were significantly more sensitive to changes in individual features (pitch or duration) than to changes in how those features were combined.

Second, the proportion of false match responses was lower for the both-out condition than for the duration-out and pitch-out conditions, $F(1, 68) = 194.33$, $p < .001$. This finding indicates that participants were more sensitive to differences in two dimensions (pitch and duration) than in just one dimension.

Third, the proportion of match responses was significantly greater for the match condition than for the switch condition, $F(1, 56) = 6.06$, $p < .05$. This finding illustrates that listeners retained some sensitivity to the manner in which pitch and duration were combined.

[2] We did not limit duration matches to one target tone because target sequences always involved a pattern of just two different duration values (which is typical of melodic phrases). If the duration of the probe tone matched just one target tone, then the six nonmatching target tones would have to be assigned the same duration, and the matching tone would be identifiable by participants as the only target tone with a distinct duration.

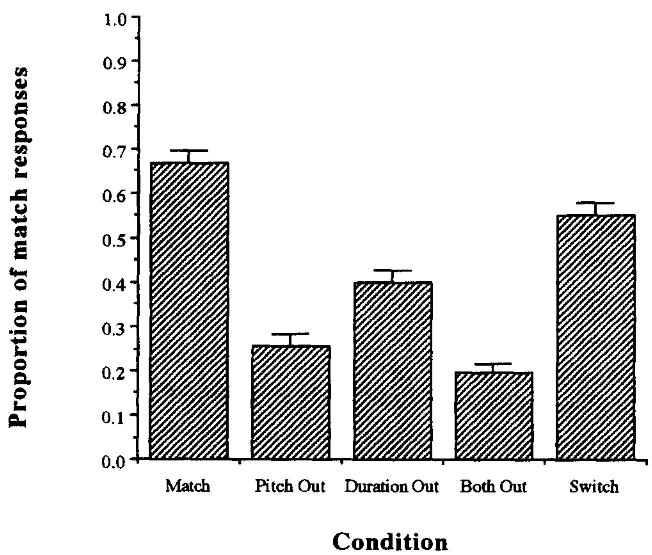

Figure 4. The proportion of errors for each of the five conditions in Experiment 2.

Signal detection analyses further confirmed that listeners had difficulty encoding combinations of pitch and duration. Sensitivity to differences between the probe tone and target tones was significantly poorer for the switch condition ($M\ d' = 0.36$) than for the pitch-out ($M\ d' = 1.24$), duration-out ($M\ d' = 0.77$), and both-out conditions ($M\ d' = 1.44$), $F(1, 51) = 44.39$, $p < .0001$. Thus, participants were significantly more sensitive to individual features in target tones than to how those features were combined. Nonetheless, the mean d' value in the switch condition was above chance, $t(17) = 2.68$, $p < .05$, indicating that listeners were not entirely insensitive to the manner in which pitch and duration were combined.

Experiment 3

The results of Experiments 1 and 2 indicate that listeners have difficulty remembering combinations of pitch and duration and implicate the occurrence of illusory conjunctions of these dimensions. Target sequences in these experiments, however, were not highly musical in that a square wave timbre was used and either there were only two tones (Experiment 1) or there were pauses between tones (Experiment 2). We conducted Experiment 3 to assess whether illusory conjunctions also occur for more naturalized melodic stimuli. Thus, sequences in Experiment 3 consisted of seven consecutive piano notes with no pauses between tones.

Method

Participants. Participants were 15 adults from the York University community, ranging in age from 19 to 55 years ($M = 25.5$). None had participated in Experiments 1 or 2. A survey revealed that participants had had between 0 and 15 years of musical training ($M = 4.8$ years). All participants reported normal hearing.

Stimuli. Each trial consisted of a seven-tone target sequence, followed by a 1.0-s pause and then a probe tone. The timbre of tones was fixed as the Piano 1 sound of the Roland Sound Canvas, which was sampled from a Steinway grand piano (for an acoustic description, see Fletcher & Rossing, 1991, 12.8). In all other respects, tones were generated in the same way as in Experiments 1 and 2. The tempo of sequences was set to 120 beats per minute, or 0.50 s per quarter note. Participants were allowed to adjust the intensity to a comfortable listening level between 60 and 70 dB SPL and heard the sequences through Sennheisser HD-480 headphones.

There were 20 target sequences. As in Experiment 2, all target sequences started and ended on the same pitch and involved an upward melodic contour for four tones followed by a downward melodic contour for three tones. The starting pitch of target sequences randomly varied between G2 and D3, and the pitch range of sequences ranged from 7 to 11 semitones. All target sequences involved rhythmic patterns of two different duration values: patterns of 500 ms and 250 ms durations, 250 ms and 125 ms durations, or 1,000 ms and 500 ms durations.

Each of the 20 target sequences was presented with each of five different probe tones (i.e., 100 trials), thus defining the five conditions described in the *Method* section of Experiment 1 (match, pitch-out, duration-out, both-out, and switch). Figure 5 provides the notation for one target sequence and the probe tones used for each of the five conditions. For conditions in which the probe tone matched a target tone in pitch, duration, or both pitch and duration, the match never occurred on the first or last target tone (to avoid enhanced performance from primacy and recency effects). For con-

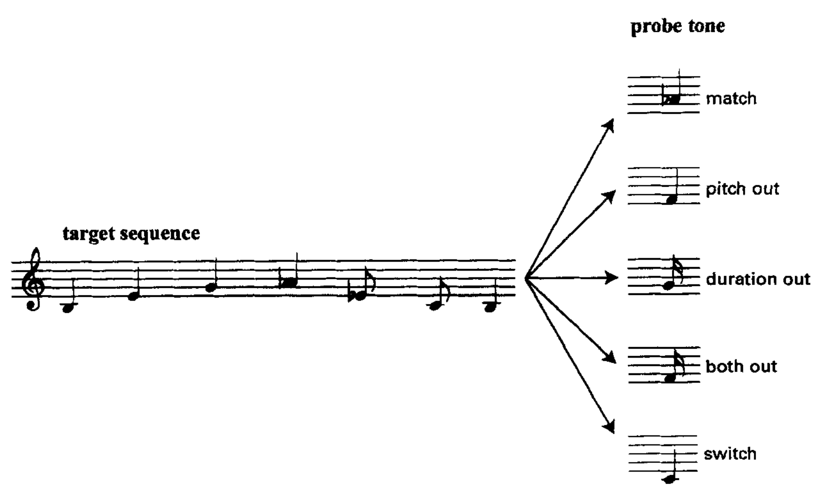

Figure 5. An illustration of a target sequence used in Experiment 3 and the probe tones used for each of the five conditions.

ditions in which the probe tone matched a target tone in duration, the match occurred with either two, three, four, or five target tones (see Footnote 2).[3]

Procedure. The procedure was identical to that used in Experiments 1 and 2.

Results and Discussion

Figure 6 displays the proportion of match responses for each of the five conditions. An ANOVA on match responses revealed a highly significant effect of condition, $F(4, 56) = 26.48, p < .001$, with a relatively high proportion of *yes* (i.e., match) responses for the match and switch conditions.

Planned comparisons confirmed two of the predictions outlined in Experiment 1. First, the proportion of false match responses was significantly lower for the duration-out and pitch-out conditions than for the switch condition, $F(1, 56) = 26.14, p < .001$. This finding indicates that listeners were significantly more sensitive to changes in individual features (pitch or duration) than to changes in how features were combined.

Second, the proportion of false match responses was lower for the both-out condition than for the duration-out and pitch-out condition, $F(1, 56) = 11.52, p < .01$. This finding indicates that participants were more sensitive to differences in two dimensions (pitch and duration) than in just one dimension.

Finally, in contrast to the results obtained using less natural melodic stimuli (Experiments 1 and 2), the proportion of false match responses for the switch condition was not significantly lower than the proportion of match responses for the match condition, $F(1, 56) = 2.44$, *ns*. This finding suggests that listeners were not sensitive to the manner in which pitch and duration were combined, and hence, responded similarly to match and switch trials.

Signal detection analyses further confirmed that listeners had difficulty encoding combinations of pitch and duration. Sensitivity to differences between the probe tone and target tones was significantly poorer for the switch condition ($M\ d' = 0.15$) than for the pitch-out ($M\ d' = 74.00$), duration-out ($M\ d' = 0.73$), and both-out conditions ($M\ d' = 1.15$), $F(1, 51) = 44.85, p < .0001$. Moreover, whereas sensitivity to differences between the probe tone and target tones was significantly greater than zero for the pitch-out condition ($M\ d' = 0.74$), $t(14) = 7.49, p < .01$; duration-out condition ($M\ d' = 0.73$), $t(14) = 5.02, p < .01$; and both-out condition ($M\ d' = 1.15$), $t(14) = 6.50, p < .01$, sensitivity was not significantly greater than zero for the switch condition ($M\ d' = 0.15$), $t(14) = 1.02$, *ns*. Thus, participants were sensitive to the individual features in target tones but insensitive to how those features were combined.

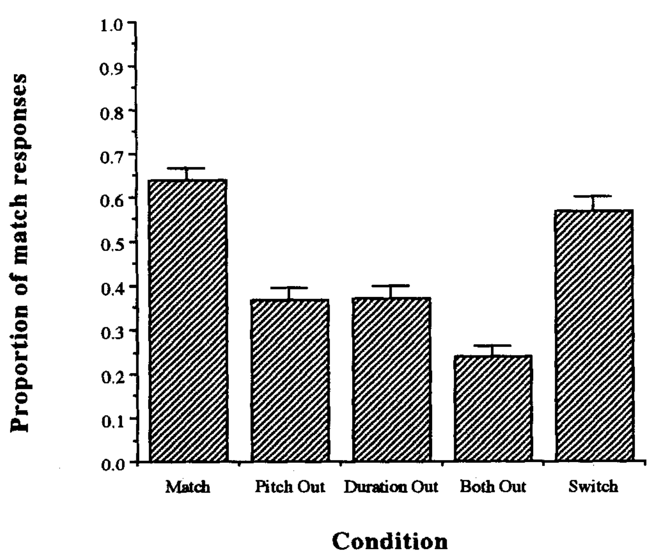

Figure 6. The proportion of errors for each of the five conditions in Experiment 3.

Model of Illusory Conjunction Rate

In all three experiments, the high error rate for the switch condition, in comparison with the relatively low error rate for other conditions, implicates the occurrence of illusory conjunctions of pitch and duration in short-term memory. A match response on the switch condition indicates that a participant falsely remembered the pitch and duration of the probe tone occurring in a single target tone, whereas, in fact, the features occurred in separate target tones. The contribution of illusory conjunctions to errors on the switch condition can be roughly estimated by comparing the error rate for the switch condition with the error rates for the pitch-out and duration-out conditions. To obtain a more precise estimate of the rate of illusory conjunctions, however, we developed a mathematical model of response errors.

The model is comparable to models of illusory conjunction rate for vision proposed by Ashby et al. (1996), insofar as we assume that response errors reflect either the misperception of a feature (a feature error) or the illusory conjunction of features (see also Batchelder & Riefer, 1999). The probabilities of three types of encoding errors are included in the model: (a) false negative feature error (fn), the probability of incorrectly encoding a feature in the target sequence, such that the participant fails to register an existing match between a feature of a target tone and a feature of the probe tone; (b) false positive feature error (fp), the probability of incorrectly encoding a feature in the target sequence, such that the participant falsely registers a match between a feature of a target tone and a feature of the probe tone; and (c) illusory conjunction (IC), the probability of incorrectly encoding the manner in which a pitch and a duration are combined.

For each condition, a response error may result from one of several encoding errors. To illustrate how different combinations of events can lead to a particular response, Figure 7 displays four of the possible encoding paths for a target tone that matches the probe tone and depicts how feature registration errors can lead to an incorrect response even when features are correctly bound (1 − IC). In this case, three of the four paths shown lead to an incorrect response. Moving from left to right, in the first path of Figure 7, the duration and pitch of the matching tone are both incorrectly encoded (false-negative duration, or fnd, and false-negative pitch, or fnp). In the second path, the duration of the matching tone is

[3] For the match and pitch-out conditions, the average number of target tones that matched the duration of the probe tone was 3.6; for the switch condition, the average number of target tones that matched the duration of the probe tone was 3.7.

incorrectly processed (fnd), whereas the pitch is correctly encoded (1 − fnp). In the third path, the duration of the matching tone is correctly encoded (1 − fnd), whereas the pitch is incorrectly encoded (fnp). In the fourth path, both the pitch and the duration of the matching tone are correctly encoded (1 − fnp and 1 − fnd). Only the latter path results in a correct response. Because the three paths that lead to a response error are mutually exclusive, their contributions to the error rate are added. The final equation used for the match condition also includes the contribution of binding errors and other conditions.

The full modeling equations and their derivations are provided in the Appendix.[4] Equations A1 through A5, respectively, describe the error paths that lead to a response error for each of the five conditions. Briefly, the equations express the rate of response errors on each condition as a function of errors in encoding tones in the target sequence; errors in encoding the probe tone were excluded on the basis that the probe tone was presented immediately prior to the response, and its features should have been highly salient in memory. We also assumed that each registered feature may be represented in only one conjunction (Treisman & Schmidt, 1982). Finally, we excluded the possibility of illusory conjunctions of incorrectly registered features, on the basis that the probability of such an error combination is likely to approach zero and would render the model unparsimonious. For example, we excluded the possibility of errors in the switch condition from the joint occurrence of (a) falsely detecting a match between the pitch of a target tone and the pitch of the probe tone (false-positive pitch, or fpp), (b) falsely detecting a match between the duration of a different target tone and the duration of the probe tone (false-positive duration, or fpd), and (c) forming an IC between these two incorrectly registered features.

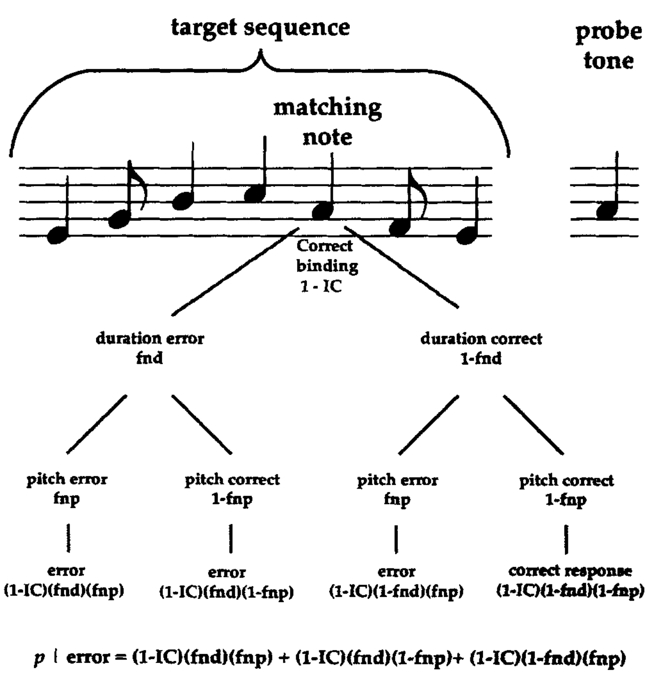

Figure 7. Four encoding paths for a target tone that matches the probe tone in the match condition of Experiment 2. IC = illusory conjunction; fnd = false-negative for duration; fnp = false-negative for pitch.

Table 1
Mean Estimates of Encoding Errors for the Three Experiments

Experiment	fnp	fnd	fpp	fpd	IC
1	.14 (.04)	.13 (.05)	.43 (.05)	.35 (.05)	.11 (.03)
2	.11 (.03)	.30 (.06)	.12 (.02)	.38 (.03)	.19 (.03)
3	.30 (.07)	.52 (.06)	.28 (.06)	.37 (.07)	.31 (.07)
Grand *M*	.18	.32	.28	.37	.20

Note. Standard errors are shown in parentheses. fnp = false-negative pitch; fnd = false-negative duration; fpp = false-positive pitch; fpd = false-positive duration; IC = illusory conjunction.

For each participant, the probability of a response error in each condition was substituted into the corresponding equation. The five parameters (encoding errors) in the equations then were estimated using Newton-Raphson's method of nonlinear minimization. This iterative procedure provides a least-squares estimate of the probability of each type of encoding error (for a related approach, see Ashby et al., 1996). Table 1 shows the means and standard errors of these estimated probabilities for each experiment. For all three experiments, the estimated probability of an illusory conjunction was significantly greater than zero ($p < .001$). Estimates of illusory conjunction rates varied across the three experiments, with the lowest estimates for Experiment 1 (two tone sequences, square wave tones) and the highest estimates for Experiment 3 (seven tone sequences, piano tones). These differences among IC estimates parallel differences in sensitivity values observed for the switch condition in the three experiments. Finally, Table 1 also illustrates that duration errors were generally more probable than pitch errors and that fp responses were more probable than fn responses.

Predicted error rates for the five experimental conditions based on these estimated probabilities are extremely accurate. Table 2 compares predicted and actual error rates for each condition, averaged across participants. For each experiment, there is a high degree of correspondence. Predicted error rates were also accurate for individual participants. As displayed in Table 3, the mean absolute difference between predicted and actual error rates by participants is only .036. These results indicate that the estimated probabilities of encoding errors provide an accurate account of response errors.

Because the modeling equations are nonlinear and solutions were limited to values between 0.0 and 1.0, the minimization

[4] The current model is reported because it is parsimonious and yielded conservative estimates of IC. Models with other assumptions were explored, but none resulted in significantly smaller estimates of IC. In one approach, we examined a more complete model that accounted for all possible errors in remembering the temporal location of a registered feature, yielding a large number of possible error combinations. This model resulted in higher estimates of IC for all three experiments. In another approach, applicable to Experiment 3, we used separate equations for conditions in which 2, 3, 4, and 5 target tones matched the probe tone in duration. For each of these conditions, we obtained estimates for each of the five error types for each participant. There was no significant difference in the estimates of IC rate depending on the number of target tones matching the probe tone in duration (i.e., 2, 3, 4, or 5), and the average estimate of IC was similar to that obtained using the presented model.

Table 2
Predicted and Actual Error Rates for the Five Experimental Conditions: Comparison of Means

Probe condition	Experiment 1		Experiment 2		Experiment 3	
	Predicted	Actual	Predicted	Actual	Predicted	Actual
Match	.26	.25	.33	.33	.34	.36
Pitch-out	.41	.41	.25	.26	.35	.37
Duration-out	.33	.35	.40	.40	.35	.37
Both-out	.20	.17	.19	.19	.27	.24
Switch	.63	.62	.54	.55	.57	.57

procedure does not guarantee accurate predictions. To illustrate, we reassigned operators in the model (addition, subtraction, multiplication, etc.) at random and repeated the estimation procedure. As expected, predictions based on this randomized model were highly inaccurate, even though this model also involved five free parameters: The average difference between predicted and actual errors, across experiments, was .41, indicating a very poor fit with the data. We conducted a repeated-measures ANOVA to compare residuals for the proposed model with residuals for the randomized model. Predictions based on the randomized model were significantly less accurate than predictions based on the proposed model for all three experiments (for all experiments, $F > 100.00$, $p < .00001$).

Figure 8 depicts how illusory conjunctions account for a proportion of the errors in the switch condition of Experiment 2. For purposes of illustration, mean estimates for Experiment 2 are used (see Table 1). As shown in the figure, each of the two paths involving an illusory conjunction is associated with an overall probability of .12. Because these paths involve the same target pitch, they are mutually exclusive. Thus, the estimated probability of paths involving illusory conjunctions is .24, which accounts for the increased error rate for the switch condition relative to other conditions. Using estimates for individual participants, the combined contribution of all paths (either involving IC or not) yield an average predicted error rate of .54, which is extremely close to the average error rate obtained for the switch condition ($p = .55$, see Table 2, Experiment 2, switch condition).

To assess the contribution of illusory conjunctions in the model, we conducted the estimation procedure once more but with the estimated illusory conjunction rate restricted to a maximum of 0.002. This step effectively removed the contribution of illusory conjunctions, yielding a more parsimonious model that includes only errors in encoding features (no-IC model). Predictions based

Table 3
Comparison of Mean Residuals for Full and No-IC Models: Experiments 1–3

Experiment	Full	No-IC	*F*	*p*
1	.039	.058	13.38	.0017
2	.039	.076	34.71	.0000
3	.030	.069	14.76	.0018
Grand *M*	.036	.068		

Note. IC = illusory conjunction.

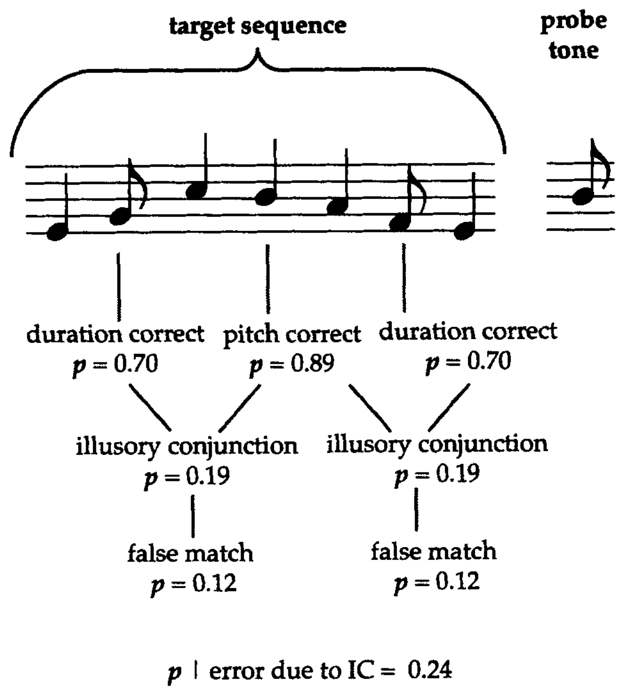

Figure 8. Error paths involving an illusory conjunction (IC) in the switch condition of Experiment 2 and the probability of this outcome based on the mean estimate of IC.

on the no-IC model, however, were less accurate. The average difference between predicted and actual errors, across experiments, participants, and conditions was .068. We conducted a repeated-measures ANOVA to compare the residuals for models with and without the contribution of illusory conjunctions. As shown in Table 2, for all three experiments, predictions based on the no-IC model were significantly less accurate than predictions based on the model that includes illusory conjunction errors. This finding further supports the occurrence of illusory conjunctions in short-term memory for target sequences.

General Discussion

The results of this investigation illustrate that illusory conjunctions of pitch and duration frequently occur in short-term memory for unfamiliar tone sequences. When a probe tone matched the pitch of one target tone and the duration of a different target tone, participants responded on well over half of the trials that the probe tone actually matched a single target tone in both pitch and duration. The occurrence of illusory conjunctions is consistent with a two-stage process of encoding unfamiliar melodies. In the first stage, melodic features such as pitch and duration are analyzed separately. Following this analysis is an integration stage in which these features are combined.

Neuropsychological evidence has confirmed that pitch and duration are neurally separated at certain stages of processing. For example, among brain-damaged patients who display impairment in singing, some have problems with pitch, while displaying normal rhythmic skills; others have problems with rhythm, while displaying normal pitch skills (Dorgeuille, 1966; cf. Peretz, 1993).

Such observations are consistent with the hypothesis that pitch and temporal information are associated with distinct brain regions (for a review, see Peretz, 1993). A similar dissociation between pitch and rhythm has been demonstrated in perceptual tasks (Peretz, 1990).

Although the occurrence of illusory conjunctions can be deduced by comparing the error rate for the switch condition with error rates for the pitch-out and duration-out condition, mathematical modeling allowed a more precise estimate of the illusory conjunction rate. The modeling acknowledges that errors on any condition can arise both from improper encoding of features (i.e., pitch and duration) and/or from illusory conjunction errors. The model illustrates how processing paths involving illusory conjunctions account for the increased error rate on the switch condition, relative to other conditions.

The estimated probability of an illusory conjunction error varied depending on the melodic context: The estimated illusory conjunction rate was lowest for Experiment 1 (.11), intermediate for Experiment 2 (.19), and relatively high for Experiment 3 (.31). Such differences imply that the properties of melodic materials affect sensitivity to combinations of pitch and duration. Experiment 1 involved just two target tones, leaving little room for difficulties in feature integration. Experiment 2 involved seven target tones, increasing the potential for confusion in feature integration. Experiment 3 also involved seven target tones, but sequences were more naturalized melodic phrases, involving piano timbres and no pauses between target tones. Interestingly, the estimated probabilities of both feature errors and illusory conjunction errors were relatively high for these naturalized melodic stimuli.

One explanation of these illusory conjunctions is that features were weakly or improperly integrated at the perceptual stage at which pitch and duration are initially conjoined. Another explanation is that connections between pitch and duration were initially strong but faded as tones were held in memory. Although further work is needed to evaluate these possibilities, it is probable that demands on short-term memory played an important role in the occurrence of illusory conjunctions. For illusory conjunctions to occur between the features of sequential tones, memory for one or both features must be involved. Moreover, because the probe tone was presented at the end of each target sequence, listeners were required to compare the properties of the probe tone with their memory for tones in the target sequence.

Although demands on short-term memory may account for the results of the current experiments, other evidence suggests that insensitivity to combinations of pitch and duration also occurs when there is little demand on memory. Thompson (1994, Experiments 1–3) found that listeners had difficulty detecting changes in the way pitch and duration were combined in melodic textures—sequences consisting of two alternating tones, each with a specific pitch and duration, repeated in a random order. This detection task should be less susceptible to memory confusions than the probe tone technique because detection of a change in the combination of pitch and duration across tones only required that the listener compare currently available features with the memories for the combination of previously available features that was repeated numerous times on a given trial.

The incidence of illusory conjunctions between pitch and duration may depend on a number of factors. For example, the results of the mathematical modeling suggest that illusory conjunctions were less probable when sequences involved two tones (Experiment 1) than when they involved seven tones (Experiments 2 and 3). A second factor is feature similarity: Tones whose pitches are psychologically related may be more likely to switch features in memory than tones whose pitches are psychological unrelated (for a discussion of the psychological similarity between pitches, see Krumhansl, 1990). Ongoing efforts in our laboratories are addressing this possibility. Third, training and familiarity may influence the likelihood of illusory conjunctions. Illusory conjunctions may be less likely for familiar melodies because of top–down influences. Moreover, training in music may help listeners to develop strategies for remembering combinations of features, even for unfamiliar musical materials. Fourth, illusory conjunctions may be more likely if the conjoined features occur in temporally adjacent tones than if they occur in temporally nonadjacent tones. That is, illusory conjunctions may result from a difficulty in retaining the temporal location of features, just as uncertainty about spatial location may give rise to illusory conjunctions of visual features (Ashby et al., 1996).

Finally, focused attention may decrease the likelihood of illusory conjunctions (Thompson, 1994; Treisman, 1986). Thompson (1994) found that distracted listeners were far less sensitive than attentive listeners to combinations of pitch and duration. It is notable, in this context, that normal listening conditions often do not involve the degree of focused attention that is typical of participants in an experimental task. Thus, the rate of illusory conjunctions under normal listening conditions may be higher than that reported here.

To conclude, the current results suggest that listeners are relatively insensitive to the manner in which pitch and duration are combined in novel melodies and that illusory conjunctions between pitch and duration frequently occur. Although listeners may be sensitive to combinations of musical dimensions for familiar materials, dimensions may be relatively dissociated for novel materials. This dissociation between musical dimensions may, in turn, have important aesthetic implications, influencing our perceptions of theme and variation, and, more generally, melodic similarity.

References

Ashby, F. G., Prinzmetal, W., Ivry, R., & Maddox, W. T. (1996). A formal theory of feature binding in object perception. *Psychological Review, 103,* 165–192.

Batchelder, W. H., & Riefer, D. M. (1999). Theoretical and empirical review of multinomial process tree modeling. *Psychonomic Bulletin & Review, 6,* 57–66.

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

Cutting, J. (1976). Auditory and linguistic processes in speech perception: Inferences from six fusions in dichotic listening. *Psychological Review, 83,* 114–140.

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.

Deutsch, D. (1986). Auditory pattern recognition. In K. Boff, L. Kaufman, & J. Thomas (Eds.), *Handbook of perception and human performance:*

Vol. 2. Cognitive processes and performance (ch. 32, pp. 1–44). New York: Wiley and Sons.

Deutsch, D. (1999). *The psychology of music* (2nd ed.). New York: Academic Press.

Dorgeuille, C. (1966). *Introduction à l'étude des amusies*. Unpublished doctoral thesis, Université de la Sorbonne, Paris.

Duncan, J., & Humphreys, G. W. (1989). Visual search and stimulus similarity. *Psychological Review, 96,* 433–458.

Efron, R., & Yund, E. W. (1974). Dichotic competition of simultaneous tone bursts of different frequency: I. Dissociation of pitch from lateralization and loudness. *Neuropsychologia, 12,* 149–156.

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

Forte, A. (1974). *Tonal harmony in concept and practice*. New York: Holt, Rinehart & Winston.

Forte, A., & Gilbert, S. (1982). *Introduction to Schenkerian analysis*. New York: WW Norton.

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

Hall, M. D., Pastore, R. E., Acker, B. E., & Huang, W. (2000). Evidence for auditory feature integration with spatially distributed items. *Perception & Psychophysics, 62,* 1243–1257.

Hindemith, P. (1946). *Elementary training for musicians*. London: Schott.

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.

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

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

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

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

Macmillan, N. A., & Creelman, C. D. (1991). *Detection theory: A user's guide*. Cambridge: Cambridge University Press.

Maddox, W. T., Prinzmetal, W., Ivry, R., & Ashby, F. G. (1994). A probabilistic multidimensional model of location discrimination. *Psychological Research, 56,* 66–77.

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.

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

Palmer, C., & Krumhansl, C. L. (1987a). Independent temporal and pitch structures in determination 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.

Peretz, I. (1990). Processing of local and global musical information in unilateral brain-damaged patients. *Brain, 113,* 1185–1205.

Peretz, I. (1993). Auditory agnosia: A functional analysis. In S. McAdams & E. Bigand (Eds.), *Thinking in sound: The cognitive psychology of human audition* (pp. 199–230). Oxford: Oxford University Press.

Peretz, I., & Kolinsky, R. (1993). Boundaries of separability between melody and rhythm in music discrimination: A neuropsychological perspective. *Quarterly Journal of Experimental Psychology: Human Experimental Psychology, 46A*(2), 301–325.

Peretz, I., & Morais, J. (1989). Music and modularity. *Contemporary Music Review, 4,* 279–293.

Prinzmetal, W., Henderson, D., & Ivry, R. (1995). Loosening the constraints on illusory conjunctions: The role of exposure duration and attention. *Journal of Experimental Psychology: Human Perception and Performance, 21,* 1362–1375.

Thompson, W. F. (1994). Sensitivity to combinations of musical parameters: Pitch with duration, and pitch pattern with durational pattern. *Perception & Psychophysics, 56,* 363–374.

Treisman, A. (1982). Perceptual grouping and attention in visual search for features and for objects. *Journal of Experimental Psychology: Human Perception and Performance, 8,* 194–214.

Treisman, A. (1986). Properties, parts, and objects. In K. Boff, L. Kaufman, & J. Thomas (Eds.), *Handbook of perception and human performance: Vol. 2. Cognitive processes and performance* (ch. 35, pp. 1–70). New York: Wiley & Sons.

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

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

Treisman, A., & Gormican, S. (1988). Feature analysis in early vision: Evidence from search asymmetries. *Psychological Review, 95,* 15–48.

Treisman, A., & Schmidt, N. (1982). Illusory conjunction in the perception of objects. *Cognitive Psychology, 14,* 107–141.

Treisman, A., & Souther, J. (1986). Search asymmetry: A diagnostic for preattentive processing of separable features. *Journal of Experimental Psychology: General, 114,* 285–310.

Wolfe, J. M., Cave, K. R., & Franzel, S. L. (1989). Guided search: An alternative to the feature integration model for visual search. *Journal of Experimental Psychology: Human Perception and Performance, 15,* 419–433.

Wolford, G. (1975). Perturbation model for letter identification. *Psychological Review, 82,* 184–199.

Appendix

Mathematical Model

The equation for each condition represents the set of encoding errors that lead to a response error. Each condition allows for a different pattern of error paths. If two errors, A and B, are mutually exclusive, their combined probability is given by $p(\text{A}) + p(\text{B})$. If two errors, A and B, are independent, their combined probability is given by $1 - (1 - p(\text{A}))(1 - p(\text{B}))$.

In the equations presented in this Appendix, fnp = false-negative for pitch; fnd = false-negative for duration; fpp = false-positive for pitch; fpd = false-positive for duration; n = number of target tones; k = number of target tones with a duration matching that of the probe tone. For Experiment 1, $n = 2$ and $k = 1$. For Experiment 2, $n = 7$ and $k = 2$. For Experiment 3, $n = 7$ and k is taken as the average number of target tones with a duration matching that of the probe tone (for the match and pitch-out conditions, $n = 3.6$; for the switch condition, $n = 3.7$).

Match Condition

Four mutually exclusive paths, representing combinations of encoded events, lead to a response error in the match condition: (1) incorrectly encoding both the matching pitch (fnp) and the matching duration (fnd), (2) correctly encoding the matching pitch (1 − fnp) but incorrectly encoding the matching duration (fnd), (3) correctly encoding the matching duration (1 − fnd) but incorrectly encoding the matching pitch (fnp), and (4) correctly encoding the matching pitch (1 − fnp) and the matching duration (1 − fnd) but failing to integrate those features (IC). Because the model excludes illusory conjunctions of incorrectly registered features, registered features in Paths 1–3 are bound together. The combined contribution of these error paths is given by their sum, or $(1 - \text{fnp})\text{fnd}(1 - \text{IC}) + (1 - \text{fnd})\text{fnp}(1 - \text{IC}) + \text{fnp}(\text{fnd})(1 - \text{IC}) + (1 - \text{fnp})(1 - \text{fnd})\text{IC}$.

The above four error paths only lead to a response error on the condition that no false matches (false positive, or fp) occur with nonmatching tones. There were k tones in the target sequence with a duration that matched that of the probe tone. One of these tones matched the probe tone in both duration and pitch, and $k - 1$ target tones matched the probe tone in duration but not pitch. A false match of the probe tone with the latter type of tone is assumed to result from correctly encoding the matching duration (1 − fnd), incorrectly encoding the pitch as matching that of the probe tone, and binding these features together (1 − IC). The joint probability of these events is given by $(1 - \text{fnd})(\text{fpp})(1 - \text{IC})$. Thus, the probability that no false matches will occur for any of these target tones is given by $(1 - (1 - \text{fnd})(\text{fpp})(1 - \text{IC}))^{k-1}$.

An average of $n - k$ target tones differed from the probe tone in both pitch and duration. A false match of the probe with one of these tones is assumed to result from incorrectly encoding both the duration and pitch of a tone as matching that of the probe tone (fpd and fpp, respectively) and correctly binding these features (1 − IC). The joint probability of these events is given by their product, or $(\text{fpd})(\text{fpp})(1 - \text{IC})$. Thus, the probability that no false matches with the probe will occur for any of these tones is given by $(1 - (\text{fpd})(\text{fpp})(1 - \text{IC}))^{n-k}$.

The combination of all error paths in this condition is expressed in Equation A1:

$$P(\text{Error}|\text{Match}) = ((1 - \text{fnp})\text{fnd}(1 - \text{IC}) + (1 - \text{fnd})\text{fnp}(1 - \text{IC}) + \text{fnp}(\text{fnd})(1 - \text{IC}) + (1 - \text{fnp})(1 - \text{fnd})\text{IC}) \times (1 - (1 - \text{fnd})\text{fpp}(1 - \text{IC}))^{[k-1]}(1 - \text{fpd}(\text{fpp})(1 - \text{IC}))^{[n-k]}. \quad \text{(A1)}$$

Pitch-Out Condition

In the sequences used for the pitch-out condition, an average of k tones per sequence had a duration that matched that of the probe tone. A false match with one of these tones is assumed to result from correctly encoding the matching duration (1 − fnd), incorrectly encoding the pitch as matching that of the probe tone (fpp), and binding these features together (1 − IC).

An average of $n - k$ tones per sequence differed from the probe tone in both pitch and duration. A false match of the probe with one of these tones is assumed to result from incorrectly encoding both the duration and pitch of a tone as matching that of the probe tone (fpd and fpp, respectively) and correctly binding these features (1 − IC). The combination of error paths is expressed in Equation A2:

$$P(\text{Error}|\text{Pitch-Out}) = 1 - (1 - (1 - \text{fnd})\text{fpp}(1 - \text{IC}))^{[k]}(1 - \text{fpp}(\text{fpd})(1 - \text{IC}))^{[n-k]}. \quad \text{(A2)}$$

Duration-Out Condition

In the sequences used for the duration-out condition, one tone matched the probe tone in pitch but not duration. For this tone, a false match is assumed to result from correctly encoding the matching pitch (1 − fnp), incorrectly encoding the duration as matching that of the probe tone (fpd), and correctly binding these features (1 − IC). $N - 1$ other tones from the target sequence differed from the probe tone in both pitch and duration. A false match of the probe with one of these tones is assumed to result from incorrectly encoding both the duration and pitch of a tone as matching that of the probe tone (fpd and fpp, respectively) and correctly binding these features (1 − IC). The combination of error paths is expressed in Equation A3:

$$P(\text{Error}|\text{Duration-Out}) = 1 - (1 - (1 - \text{fnp})\text{fpd}(1 - \text{IC}))((1 - \text{fpp}(\text{fpd})(1 - \text{IC}))^{n-1}). \quad \text{(A3)}$$

Both-Out Condition

In the sequences used for the both-out condition, seven tones differed from the probe tone in both pitch and duration. A false match with one of these tones is assumed to result from the combination of three events: incorrectly encoding the duration as matching that of the probe tone (fpd), incorrectly encoding the pitch as matching that of the probe tone (fpp), and correctly binding these features (1 − IC). The combined influence of the seven independent opportunities for this path (one per tone) is expressed in Equation A4:

$$P(\text{Error}|\text{Both-Out}) = 1 - (1 - \text{fpp}(\text{fpd})(1 - \text{IC}))^{n}. \quad \text{(A4)}$$

Switch Condition

Four different paths, representing different combinations of errors, would result in an error response in the switch condition. These paths are described separately below.

Illusory Conjunction

One target tone matched the probe tone in pitch, and an average of k tones per sequence matched the probe tone in duration. A false match is assumed to result if the matching pitch from one tone and the matching duration from another tone are correctly encoded (1 − fnp and 1 − fnd,

(*Appendix continues*)

respectively), and an illusory conjunction is formed between those features (IC).

Pitch-Error

There were k tones per sequence with a duration that matched that of the probe tone. A false match with one of these tones is assumed to result from correctly encoding the matching duration (1 − fnd), incorrectly encoding the pitch as matching that of the probe tone (fpp), and binding the features together (1 − IC).

Duration-Error

One tone matched the probe tone in pitch but not duration. For this tone, a false match is assumed to result from correctly encoding the matching pitch (1 − fnp), incorrectly encoding the duration as matching that of the probe tone (fpd), and correctly binding these features (1 − IC).

Pitch-Duration Error

An average of $n - 1 - k$ tones per sequence differed from the probe tone in both pitch and duration (for Experiment 1, this value is 0). A false match with one of these tones is assumed to result from incorrectly encoding both the pitch and duration of a tone as matching the pitch and duration of the probe tone (fpp and fpd, respectively) and correctly binding those features (1 − IC).

Combined Error Paths

The k illusory conjunction paths are mutually exclusive of each other (because all involve the same encoded pitch). The illusory conjunction paths also are mutually exclusive of the pitch-error and duration-error path (because they involve the same encoded durations or pitch, respectively). The k pitch-error paths are independent of each other and independent of the duration-error path. Finally, the pitch/duration error paths are independent of the illusory conjunction paths, the pitch-error paths, and the duration-error path. The resulting combination of error paths is expressed in Equation A5:

$$\mathrm{P(Error|Switch)} = (\mathrm{k(IC)(1 - fnd)(1 - fnp)} + 1 - (1 - (1 - \mathrm{fnd})\mathrm{fpp}(1 - \mathrm{IC}))^{[k]}(1 - (1 - \mathrm{fnp})(\mathrm{fpd})(1 - \mathrm{IC}))) \cdot (1 - \mathrm{fpd}(\mathrm{fpp})(1 - \mathrm{IC}))^{[n-1-k]}. \quad \text{(A5)}$$

Received October 23, 1998
Revision received November 12, 1999
Accepted April 19, 2000 ■