

## Research Article

## Is Relative Pitch Specific to Pitch?

Josh H. McDermott, Andriana J. Lehr, and Andrew J. Oxenham

*University of Minnesota*


---

**ABSTRACT**—*Melodies, speech, and other stimuli that vary in pitch are processed largely in terms of the relative pitch differences between sounds. Relative representations permit recognition of pitch patterns despite variations in overall pitch level between instruments or speakers. A key component of relative pitch is the sequence of pitch increases and decreases from note to note, known as the melodic contour. Here we report that contour representations are also produced by patterns in loudness and brightness (an aspect of timbre). The representations of contours in different dimensions evidently have much in common, as contours in one dimension can be readily recognized in other dimensions. Moreover, contours in loudness and brightness are nearly as useful as pitch contours for recognizing familiar melodies that are normally conveyed via pitch. Our results indicate that relative representations via contour extraction are a general feature of the auditory system, and may have a common central locus.*

---

On hearing a melody, people typically remember little about the exact pitches of its notes, but are fairly adept at remembering the pattern of changes between notes (Attneave & Olson, 1971). Transpositions of melodies—versions shifted up or down in pitch—are therefore easily recognized, as the relationships between notes are preserved despite the changes in individual pitches. Young infants share this ability to recognize melody transpositions (Plantinga & Trainor, 2005; Trehub, Bull, & Thorpe, 1984); evidently, relative representations for pitch emerge without training or extensive auditory experience. Relative pitch is critical to recognizing music and speech prosody across the varying pitch ranges of different instruments and speakers, and is thus a signature feature of how humans process sound.

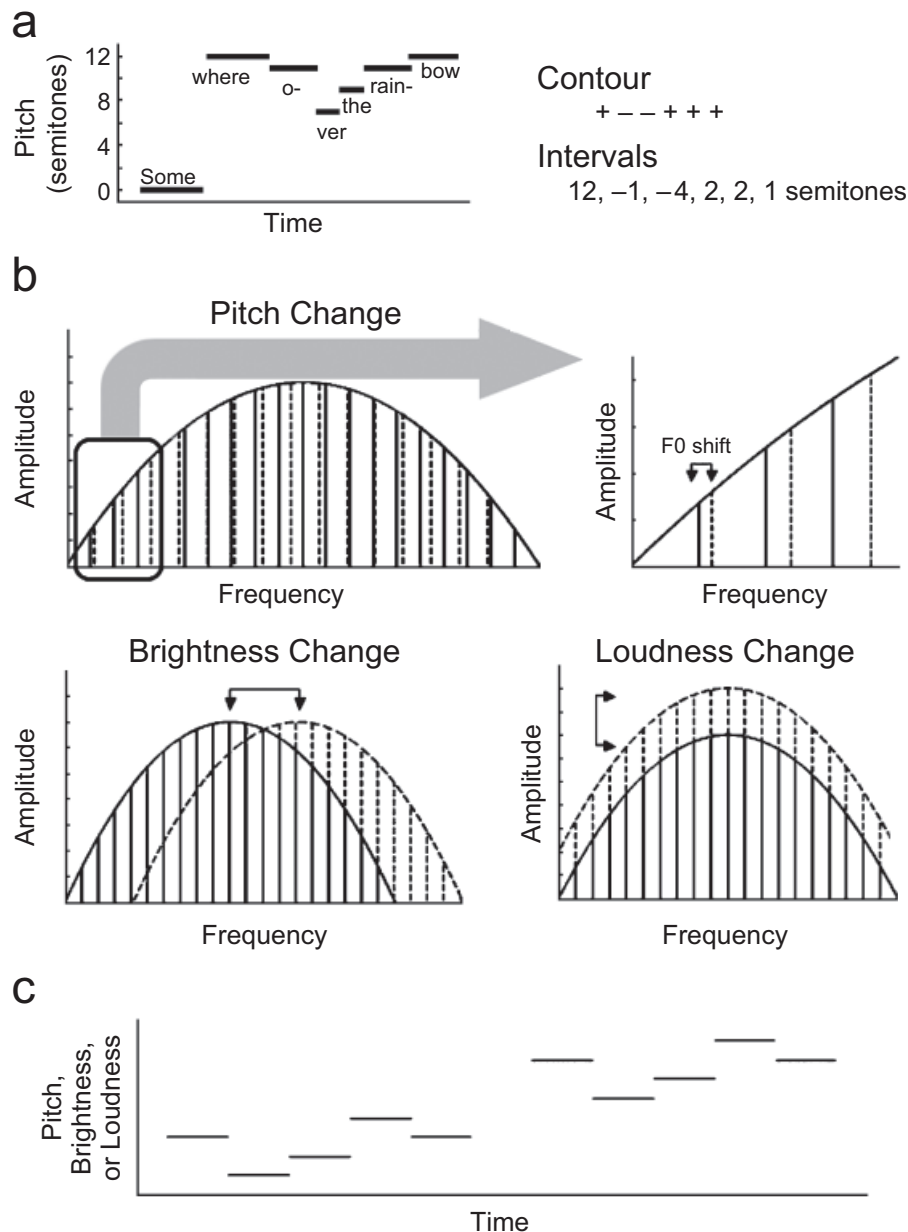
---

Address correspondence to Josh H. McDermott, Department of Psychology, University of Minnesota, N218 Elliott Hall, 75 East River Rd., Minneapolis, MN 55455-0344, e-mail: joshmcd@umn.edu.

Despite the importance of relative pitch, little is known about how it is extracted. In music cognition, it is common to distinguish between two components of relative pitch: contour and interval (Dowling & Fujitani, 1970; see Fig. 1a). The contour is the sequence of pitch-change directions from note to note; it contains the sign of each change, but not its magnitude. The intervals are the exact amounts by which the pitch changes from note to note. For randomly generated novel melodies, recognition is dominated by the contour (Dowling, 1978; Dowling & Fujitani, 1970). For familiar melodies, both contour and intervals are critical to recognition. Interval sizes must be produced correctly in order for recognition to occur with consistency, though the contour alone can produce some degree of recognition (Dowling & Fujitani, 1970), and has been recognized as important by many music theorists (Morris, 1993).

To gain insight into the mechanisms of relative pitch, we tested whether listeners could use comparable relative representations for other dimensions of sound. Our experiments were motivated by the question of what provides pitch (as opposed to any other auditory dimension) its special status in music. Melodic structure in every culture known to Western scholars is determined by pitch variation (Patel, 2008), even though musical patterns could in principle be generated by other means, such as variation in timbre (a fact that has not escaped the notice of some modern composers; Marvin, 1995). We wondered whether relative representations might be unique to pitch, or whether they might be generated for other dimensions as well.

We focused on two dimensions well characterized by prior research: loudness and brightness. Brightness is one of the most salient dimensions of timbre (McAdams, Winsberg, Donnadieu, De Soete, & Krimphoff, 1995). It is the perceptual correlate of the center of mass of the frequency spectrum (Fig. 1b, lower left). Sounds with more high-frequency energy are brighter than those with less high-frequency energy. Brightness varies across vowels and instruments, and can be varied orthogonally to pitch and loudness (Fig. 1b). Variations in both loudness and brightness are common in everyday stimuli. Rapid changes in the shape of the spectrum (reflected by formant positions) characterize sequences of vowels, and local changes in intensity often signal



**Fig. 1.** Illustrations of (a) contour and intervals in a musical melody; (b) manipulations of pitch, brightness, and loudness in a complex tone; and (c) the contours presented in Experiment 1. The diagram in (a) shows the first few notes of “Somewhere Over the Rainbow,” with contour and intervals specified. Experiments 1 through 3 measured the ability to recognize contours in pitch, brightness, and loudness. In the diagrams illustrating these manipulations (b), straight lines denote frequency components; the curved line is the spectral envelope. Solid and dashed lines denote two different sounds. Arrows indicate the stimulus change producing each perceptual change. Pitch-varying stimuli were generated by changing the fundamental frequency (F0) of complex tones, keeping the spectral envelope fixed; brightness-varying stimuli were generated by shifting the spectral envelopes of complex tones, keeping the F0 fixed at 100 Hz; and loudness-varying stimuli were generated by altering the level of a burst of broadband noise. See the text for more details. In Experiments 1 through 3, participants heard two “melodies” (contours of pitch, loudness, or brightness) on each trial, and their task was to indicate whether the two melodies were the same or different. In the example illustrated here (c), the two contours are the same.

stresses and accents in speech (Kochanski, Grabe, Coleman, & Rosner, 2005). Patterns in dimensions other than pitch are thus not without ecological validity, but it remained to be seen whether such variations produce similar representations.

With these issues in mind, we modified standard melody-recognition tasks to test whether relative representations could be generated via dimensions other than pitch. Our first experiment demonstrates that subjects can recognize transpositions of

contours in loudness and brightness. The second and third experiments show that contours in one dimension can be recognized in another dimension, suggesting that similar representations are used for different dimensions. The final experiment provides further support for this notion by showing that contours in brightness and loudness can support the recognition of familiar songs. Our results indicate that contour representations may be a general feature of the auditory system.

## GENERAL METHOD

We began our study with the classic paradigm used by Dowling and his colleagues (Dowling, 1978; Dowling & Fujitani, 1970) to demonstrate the importance of the melodic contour. They presented subjects with a randomly generated five-note melody, followed by a second five-note melody that was transposed up or down (Fig. 1c). The contour and interval composition of the two melodies were varied separately, and each of these components could be either the same or altered in the second melody. Dowling and his colleagues found that recognition depended primarily on the contour.

For our first three experiments, we extended Dowling's paradigm to allow stimuli to vary in pitch, brightness, or loudness. The task was always to judge whether the pattern of variation (contour) in the two stimuli was the same or different. On half the trials in each condition, the two stimuli had the same contour; on the other half, they had different contours (i.e., the second melody was regenerated such that at least one of the intervals differed in sign from the first melody).

Following each trial, subjects clicked a button to select one of four responses ("sure different," "different," "same," "sure same"), depending on their level of confidence. Subjects were instructed to attempt to use all four responses equally often. Hits and false alarms were computed on the basis of whether the contour was in fact the same or different. These quantities were converted into a receiver-operating characteristic (ROC) curve for each condition, and the area underneath the curve was used as the measure of performance. This area always lies between 0 and 1; .5 corresponds to chance performance. The ROC area is equal to the proportion correct of an equivalent unbiased observer in a dual-pair comparison (4IAX) task (MacMillan, Kaplan, & Creelman, 1977; Micheyl & Dai, in press), a task that would be less intuitive and slower to administer. Because we were interested in what subjects would naturally perceive, rather than in what they could be trained to perceive, no feedback was given. Stimuli were delivered to subjects in a soundproof booth via headphones. Unless otherwise noted, statistical significance was assessed with one-sample *t* tests comparing an ROC area to .5.

## EXPERIMENT 1: CONTOURS IN DIMENSIONS OTHER THAN PITCH

In Experiment 1, we tested whether subjects could recognize "transpositions" in dimensions other than pitch. On each trial,

subjects were presented with two "melodies"—sequences of five "notes" that varied along different ranges of the same dimension (pitch, brightness, or loudness; see Fig. 1c). These three dimension conditions were crossed with two interval conditions. Specifically, when the first and second contours were the same, the second stimulus could either be an exact transposition of the first (same-interval condition) or have the same contour but altered intervals (altered-interval condition). In this latter case, the second melody had intervals that were the same sign as those of the first melody, but that always differed in magnitude. Subjects completed three blocks of 10 trials in each of the six conditions; both conditions and trials within conditions were randomly ordered. Seventeen undergraduates participated (13 females, 4 males; mean age = 21.82 years, *SE* = 1.24; average of 6.2 years of music training, *SE* = 1.34).

### Stimulus Details

The melodies were generated by choosing four intervals from the set  $\{-2, -1, 1, 2$  semitones $\}$ . Each interval had an equal probability of being selected, and sampling was done with replacement. The pitch-varying stimuli were generated by changing the fundamental frequency (F0) of complex tones, keeping the spectral envelope fixed (Fig. 1b, top). The lowest note of the first melody always had an F0 of 100 Hz, and the lowest note of the second melody always had an F0 6 semitones above 100 Hz (141.4 Hz). The spectral envelope was a Gaussian function that was centered at 1000 Hz and had a standard deviation of 250 Hz on a linear amplitude scale.

The brightness-varying stimuli were also generated from complex tones, by shifting the spectral envelope while keeping the F0 fixed at 100 Hz (Fig. 1b, bottom left). To mirror the logarithmic scaling of frequency, we scaled the spectral envelope in proportion to the center frequency (standard deviation was set to 25% of the centroid), normalizing the amplitude to maintain a constant root-mean-square value. To partially compensate for differences in discrimination thresholds between dimensions, we used shifts in the spectral centroid that were twice the size of the shifts in F0 (in semitones) used in the pitch-varying conditions. The lowest notes of the first and second brightness-varying melodies always had spectral envelopes centered at 1000 Hz and 2119 Hz, respectively. Both pitch- and brightness-varying stimuli were presented at 70 dB SPL.

The loudness-varying stimuli were generated by altering the level of a burst of broadband noise (bandwidth of 1500 Hz, centered at 1000 Hz, generated by setting the coefficients of frequencies outside the spectral passband to zero). Pilot experiments yielded similar results using tones instead of noises. The first melody always ranged from 45 to 65 dB SPL, and the second melody ranged from 65 to 80 dB (the second melodies were limited to a 15-dB range to avoid uncomfortably loud levels). To maximize the discriminability of the contours, we mapped the range of the melody onto the full range of intensity on each trial.

All notes in all conditions were given the same amplitude envelope (onset and offset ramps of 100 ms) and were 300 ms in total duration. Notes within a melody were presented back-to-back, and the first and second melodies were separated by a 300-ms gap.

## Results

Although performance was slightly better for the pitch-varying stimuli than for the loudness- and brightness-varying stimuli, subjects were substantially above chance at recognizing transposed patterns in all three dimensions (see Fig. 2a)—pitch:  $t(16) = 29.57, p < .0001$ ; brightness:  $t(16) = 4.82, p < .0001$ ; loudness:  $t(16) = 18.36, p < .0001$ . Performance was slightly worse when the intervals were altered than when they remained the same,  $F(1, 16) = 11.32, p = .004, h_p^2 = .41$ , but was still far above chance for all the dimensions—pitch:  $t(16) = 19.99, p < .0001$ ; brightness:  $t(16) = 4.92, p < .0001$ ; loudness:  $t(16) = 11.23, p < .0001$ . There was no interaction between dimension and interval condition. The stimulus changes were not equated for discriminability, and any advantage for pitch could be due to the fact that just-noticeable differences (JNDs) for shifts in spectral centroid and intensity are much larger, relative to their dynamic range, than JNDs for F0. The results display the hallmarks of a contour representation—invariance across transposition, and greater sensitivity to the sign of change than to the magnitude (as evidenced by the small effect of interval condition).

### EXPERIMENT 2: RECOGNIZING PITCH CONTOURS REPLICATED IN OTHER DIMENSIONS

The apparent existence of contourlike representations for brightness and loudness raises the question of whether a pitch contour can be recognized when replicated in another dimension. To test this possibility, we ran an experiment in which the first stimulus was always defined by pitch variation, but the second stimulus could vary in pitch, brightness, or loudness. The task was to judge whether the two stimuli contained the same pattern.

Given a pitch-defined melody, there are multiple patterns in loudness or brightness that could be said to have the same contour. First, the scaling of one dimension relative to the other is not fixed—a given pitch change could be made to correspond to a loudness or brightness change of arbitrary size. Second, and perhaps less obviously, the polarity of the changes in the second dimension is also not fixed—an increase in pitch could be mapped to an increase in loudness or brightness, or to a decrease. Although it might be intuitive to map pitch increases onto loudness or brightness increases, a priori there is no reason to choose this over the reverse mapping. We therefore included two conditions with brightness and loudness melodies, one for each of these possible mappings. Because the scaling of a pattern ought not to matter much for contour extraction as long as the changes are readily perceived, we mapped each pitch mel-

ody to as large a range of the other dimension as was practical, preserving the relative size of the intervals. We also included two pitch conditions, one in which the second melody on same-contour trials had the same intervals as the first melody, and one in which the intervals were altered, as in Experiment 1.

Stimuli were generated as in the previous experiment, except that the F0 of the lowest pitch of the first melody was always 150 Hz, and the F0 of the lowest pitch of the second melody was 50 Hz. Twenty-nine undergraduates participated (19 females, 10 males; mean age = 22.63 years,  $SE = 0.82$ ; average of 4.93 years of music training,  $SE = 0.83$ ).

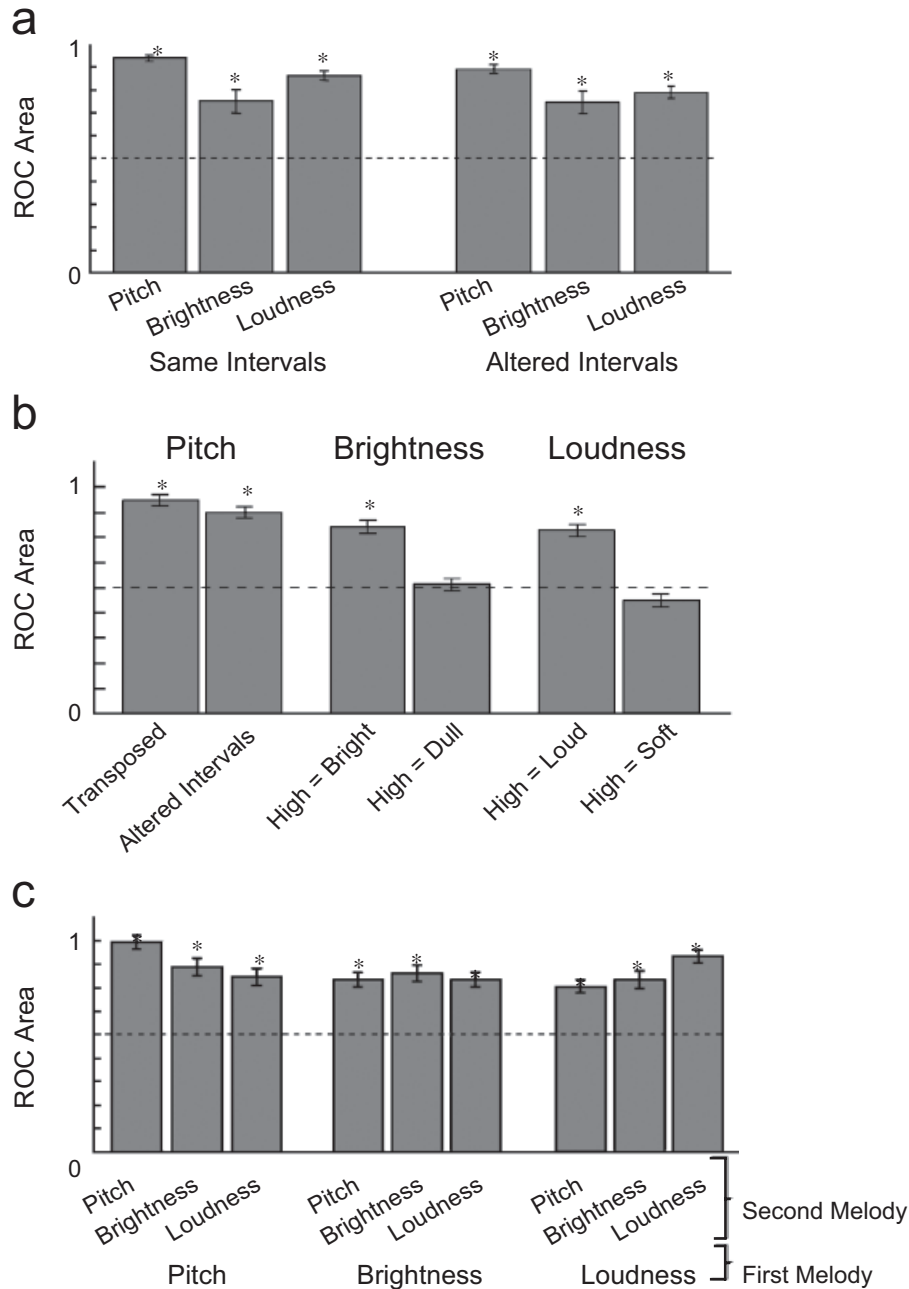
As in Experiment 1, subjects performed best when the test melody was itself defined by pitch. Performance was well above chance in these conditions,  $t(28) = 15.76, p < .0001$ , and was little impaired when the intervals were altered, remaining significantly above chance,  $t(28) = 13.0, p < .0001$  (see Fig. 2b). Performance was also high when the test melody was defined by brightness or loudness, but only when pitch increases were mapped to brightness or loudness increases—brightness:  $t(28) = 9.10, p < .0001$ ; loudness:  $t(28) = 9.82, p < .0001$ . When the reverse mapping was used, subjects did not match contours across dimensions—brightness:  $t(28) = 0.52, p = .6$ ; loudness:  $t(28) = -2.02, p = .05$ . There is evidently a natural mapping from pitch increases to brightness and loudness increases, and a contour of the opposite polarity is not heard as similar.

### EXPERIMENT 3: MATCHING CONTOURS ACROSS DIMENSIONS

We next tested whether the ability to match contours across dimensions depended on the first pattern being defined by pitch. The paradigm was identical to that of Experiment 2 except that both the first melody and the second melody could be defined by pitch, brightness, or loudness (nine conditions;  $3 \times 3$  design). The contours were all mapped in the polarity that yielded high performance in Experiment 2. Stimuli were generated as in Experiment 1. Twenty undergraduates participated (16 females, 4 males; mean age = 22.15 years,  $SE = 1.2$ ; average of 5.55 years of music training,  $SE = 1.01$ ).

As Figure 2c shows, performance was best when the matched contours were within the same dimension. This resulted in an interaction between the dimension of the first stimulus and the dimension of the second stimulus,  $F(4, 76) = 14.37, p < .0001, h_p^2 = .43$ . However, performance remained high when the dimensions of the first and second stimuli differed— $t(19)$ s ranged from 6.49 to 13.77, all significant at  $p < .0001$ . There is evidently some cost to comparing contours across dimensions, but not much.

One explanation for these results might be that subjects converted each contour to a verbal description (e.g., “up-down-down-up”), and then compared these descriptions, rather than comparing the contours themselves. The fast pace of the melodies and the short interval separating them made this a priori



**Fig. 2.** Results for Experiments 1 through 3. Each graph shows the area under the receiver-operating characteristic (ROC) curve as a function of condition. In Experiment 1 (a), participants judged whether contours in pitch, brightness, and loudness were the same or different across two presentations. On same-contour trials the intervals could be the same or different in the two presentations. In Experiment 2 (b), the first stimulus on each trial was defined by pitch variation, and the second stimulus was defined by variation in pitch, brightness, or loudness. There were two pitch conditions; the intervals of the first and second stimuli were the same in one of these conditions and different in the other. There were also two brightness conditions and two loudness conditions, across which the polarity of the changes in the second stimulus was manipulated such that increases in pitch were mapped either to increases or to decreases in loudness or brightness. In Experiment 3 (c), both the first melody and the second melody could be defined by pitch, brightness, or loudness, and all contours were mapped with the polarity that yielded high performance in Experiment 2. Dashed lines represent the level of chance performance. Error bars denote standard errors of the means; asterisks denote performance significantly different from chance,  $p < .001$ .

unlikely, but to help rule out this alternative we debriefed subjects after the experiment. We asked them what strategies they had used and specifically if they had found themselves verbalizing the contour shape and, if so, how much of the time. Only 5 of the 20 subjects reported having used such a strategy, and then only for a minority of the time (fewer than half of the trials in all cases). Verbal recoding thus seems unlikely to have played a significant role.

#### EXPERIMENT 4: RECOGNITION OF FAMILIAR MELODIES

As a further test of the equivalence of contours in different dimensions, we tested recognition of 10 well-known songs (nursery rhymes, Christmas carols, etc.) played with pitch, brightness, or loudness. Subjects were not told the melodies in advance. After hearing each stimulus, they were prompted to type in the name of the tune (they had the option of not responding). Responses were coded by a researcher who was blind to the condition.

To measure the contribution of the contour to recognition, we included conditions in which the intervals were stretched by a factor of 2, a manipulation that preserved the contour but altered the identity of its intervals. Interval stretching is less meaningful for brightness and loudness than for pitch, as it is not obvious that there is any correspondence between interval sizes across dimensions; the conditions with stretched brightness and loudness contours thus served as controls. Because the melodies were not rhythmically identical, we included a condition in which only the duration varied between notes, to assess the extent to which rhythm alone could support recognition.

The brightness melodies were given the same step sizes as the pitch melodies (the spectral envelope was shifted by the same number of semitones). The loudness melodies were generated by mapping the pitch range of a melody onto the intervals [64–82 dB SPL] (unstretched) and [46–82 dB SPL] (stretched). A quarter note was 300 ms in duration. Stimuli were otherwise generated with the same parameters as the second stimuli in Experiment 2. The experiment looped through the set of 10 melodies 10 times; each time the melodies were in random order, and assignment of melodies to conditions was varied such that each melody was played once per test-melody condition (pitch, stretched pitch, brightness, stretched brightness, loudness, stretched loudness, rhythm, plus three additional conditions not discussed here). Thirty undergraduates participated (22 females, 8 males; mean age = 21.83 years,  $SE = 0.8$ ; average of 3.3 years of music training,  $SE = 0.71$ ).

Precise intervals are known to be important for the recognition of familiar melodies, and our results confirm this, as stretching the intervals in the pitch stimuli impaired performance,  $t(29) = 6.03, p < .0001$  (paired  $t$  test; see Fig. 3). Nonetheless, we found that contour alone could partially support recognition, as performance was better in the stretched-pitch condition than in the rhythm-only condition,  $t(29) = 8.97, p < .0001$  (paired  $t$  test).

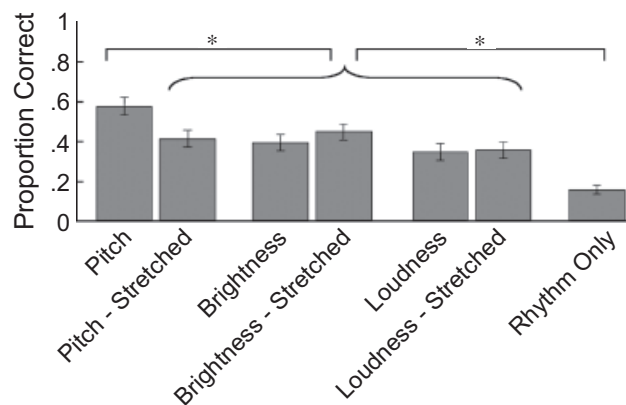


Fig. 3. Results from Experiment 4: proportion of familiar melodies correctly identified when played in pitch, brightness, or loudness, with or without stretched intervals, and when played with notes that varied only in duration (thus producing exclusively rhythmic variation). Error bars denote standard errors of the means. Asterisks indicate significant differences between conditions,  $p < .001$ .

Notably, performance with brightness and loudness contours was similar to that for stretched-pitch contours. This indicates that contours in dimensions other than pitch can also be used to recognize melodies. Post hoc comparisons (paired  $t$  tests, Bonferroni corrected) revealed that the stretched-pitch, brightness, stretched-brightness, loudness, and stretched-loudness conditions did not differ significantly from each other ( $p > .05$  in all cases), and that the main effect of condition,  $F(6, 174) = 35.188, p < .0001, h_p^2 = .55$ , was driven by the differences between those five conditions and the pitch and rhythm-only conditions (see Fig. 3). By comparison, there were no differences across conditions in the proportion of incorrect responses (in an analysis omitting nonresponses),  $F(6, 174) = 0.96, p = .46$ . Thus, criterion differences cannot explain the results.

Because the melodies were presented multiple times, we could assess the effect of foreknowledge by comparing recognition over the course of the experiment. Comparing performance for Presentations 1 through 3, 4 through 6, and 7 through 9 of each melody revealed an effect of repeated presentation,  $F(2, 58) = 14.9, p < .0001, h_p^2 = .34$  (on average, recognition improved 11.78% from Presentations 1–3 to Presentations 6–9,  $SE = 2.05\%$ ), but no interaction with condition,  $F(7.55, 218.87) = 0.49, p = .85$ . Even in the first few presentations, when listeners knew little about what melodies to expect, loudness and brightness contours were nearly as useful as pitch contours for music recognition.

In an early study of melody recognition, Moore and Rosen (1979) failed to obtain good performance when melodies were defined by loudness contours. Given the numerous differences between their study and the present one, it is not clear what accounts for this difference in results. Some potentially important factors include Moore and Rosen's use of a pure-tone (instead of a noise) carrier, the pitch of which might have interfered with the loudness contour, and their use of many pitch-based

conditions with only a single nonpitch condition, which may have primed listeners to expect pitch variations.

### PITCH-MATCHING CONTROL EXPERIMENT

One trivial explanation of our results is that changes in spectral envelope and intensity affect a sound's pitch. According to this hypothesis, subjects might recognize melodies defined by brightness and loudness merely because of the pitch contours produced by changes in these dimensions. To test this possibility, we measured the effect of these dimensions on pitch.

At the start of a trial, subjects heard a 300-ms standard stimulus, then a 300-ms gap, then a 300-ms pure tone. They could then adjust the pure tone's frequency up or down with the click of a button. Once an adjustment was made, the cycle repeated; this process continued until subjects indicated via a button click that the pitches matched. Twenty trials were run per condition; both conditions and trials within conditions were randomly ordered. There were six conditions, four testing the effect of loudness on pitch (standard stimuli were complex tones or noise bursts at 55 or 75 dB SPL) and two testing the effect of brightness on pitch (standard stimuli were complex tones with spectral centroids of 1000 Hz or 1682 Hz). The complex tones had missing F0s, drawn randomly from 50 through 100 Hz (fixed within a trial). The noise bursts were centered randomly between 700 and 1300 Hz (bandwidth = 1.5 times the center frequency). Stimuli were otherwise generated as were the notes in the melodies of Experiment 1.

The effect of intensity or brightness on pitch was measured by taking the difference between the frequency of the pure-tone match and the actual F0 of the complex tone, or the center frequency of the noise band, and comparing this difference across brightness or loudness conditions. If brightness or loudness affects perceived pitch, pitch errors would be expected to differ across the conditions. Fourteen undergraduates participated (9 females, 5 males; mean age = 23.57 years,  $SE = 1.74$ ; average of 4.6 years of music training,  $SE = 1.05$ ).

As Figure 4 shows, there was little effect of intensity on pitch matches—tone standard:  $F(1, 13) = 1.41$ ,  $p = .26$ ; noise standard:  $F(1, 13) = 0.17$ ,  $p = .69$ . Pitch errors occurred occasionally for the tones, and were common for the noise bursts, but were no more frequent for high-intensity stimuli than for low-intensity stimuli. This finding is consistent with prior work showing weak and inconsistent effects of intensity on pitch (Verschuure & van Meeteren, 1975).

There was a small influence of the spectral envelope on pitch matches,  $F(1, 13) = 6.95$ ,  $p = .02$ ,  $h_p^2 = .35$ , which was also consistent with prior work (Warrier & Zatorre, 2002). Inspection of the histograms in Figure 4 reveals that this effect was mainly due to an increase in upward octave errors when the spectral envelope was high; the majority of pitch matches in both the high-envelope condition and the low-envelope condition were

within a semitone of the true F0. It seems unlikely that this weak effect of the spectral envelope could be responsible for the robust contour discrimination in Experiments 1 through 4, though it is difficult to completely rule out a small role for the effect. Overall, however, the perception of brightness and loudness contours does not seem to be due to their effect on the pitch of individual sounds.

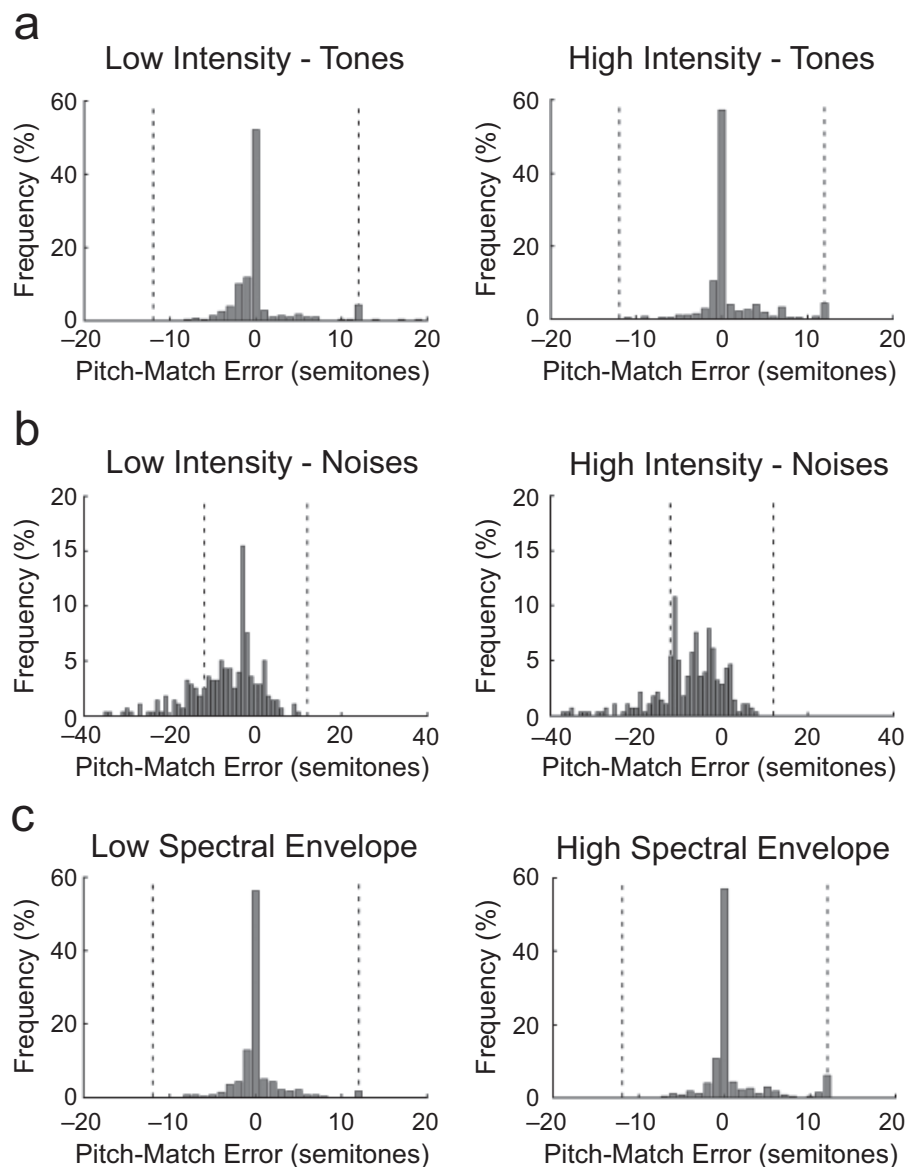
### DISCUSSION

Although relative representations have traditionally been associated with pitch, our results indicate that they may be a general feature of the auditory system. Observers were usually best at recognizing transpositions of pitch patterns, but possessed high degrees of facility for recognizing transpositions of loudness and brightness patterns as well. These results implicate representations of the changes between successive sounds, rather than of their absolute values. There were not obvious qualitative differences in performance across dimensions. Indeed, contours in one dimension could be recognized when replicated in a different dimension, which suggests that contours in different dimensions may have a common representation, or at least representations that are sufficiently similar to permit easy comparisons. This notion is underscored by our results for recognition of familiar melodies; brightness and loudness contours were recognized almost as well as pitch contours.

It remains unclear whether the similarity of contour perception across dimensions is mirrored in the perception of intervals. Fully resolving this issue will require further studies specifically targeting interval perception. Nonetheless, at present, it is at least conceivable that domain-general contour representations could exist separately from pitch-specific interval representations. This would be consistent with evidence from neuropsychology suggesting separate neural loci for contour and interval representations (Liegeois Chauval, Peretz, Babai, Laguitton, & Chauvel, 1998; Peretz & Coltheart, 2003).

Another remaining question is whether the same contour in different dimensions produces a single contour representation in the brain or two distinct but similar representations. Behavioral experiments seem unlikely to resolve this issue, but brain imaging might. It would also be interesting to explore whether amusic individuals with impairments in discrimination of pitch change (Hyde & Peretz, 2004; Semal & Demany, 2006; Stewart, von Kriegstein, Warren, & Griffiths, 2006) are impaired at hearing patterns of change in other dimensions as well.

One striking feature of our results is that contour recognition across dimensions depended on the mapping between dimensions (see Fig. 2b, brightness and loudness conditions). Brightness and loudness increases are heard as similar to pitch increases, but not as similar to pitch decreases. This effect is reminiscent of prior reports of perceptual interference among pitch, brightness, and loudness changes (Melara & Marks, 1990; Neuhoff, Kramer, & Wayand, 2002). It remains to be seen



**Fig. 4.** Results from the control experiment, in which participants matched pure tones to (a) complex tones at low and high intensity, (b) noise bursts at low and high intensity, and (c) complex tones with low and high spectral-envelope centroids. The histograms show the prevalence of matching errors (difference between the frequency of the pure-tone match and the fundamental frequency of the complex tone or the center frequency of the noise burst), measured in semitones. The histogram bins are 1 semitone wide, centered at integer numbers of semitones. The dashed lines denote one octave above and below the true fundamental frequency (or center of band for noise stimuli).

whether this dependence is a function of learned associations (Neuhoff, McBeath, & Wanzie, 1999; e.g., being learned from covariation among pitch, brightness, and loudness in speech and music), or whether it relates to the vertical metaphor used to describe all three dimensions in many languages.

If relative representations are not unique to pitch, what explains pitch's central role in music? Comparably precise interval information may not be available for other dimensions of sound, and this lack of precise information may limit their role in musical structure. However, another striking property of pitch

perception becomes apparent when one attempts to construct melodies in other dimensions: Discriminability is markedly better for pitch than for other dimensions. A semitone change is well above threshold for normal listeners, but is tiny relative to the range over which melodic pitch can be heard (approximately 80 semitones). Comparable step sizes relative to thresholds in loudness or brightness, by comparison, would yield a small fraction of this range (Gagne & Zurek, 1988; Viemeister & Bacon, 1988), and in many other dimensions the number of distinguishable steps would be even smaller. As a result, pitch



variation appears to have much greater expressive capacity than other auditory dimensions, as it provides much more room for melodies to vary. Thus, although the mechanisms for representing contours are not specific to pitch, the ability to use small changes to construct such contours might be, and could help explain the ubiquity of pitch in music.

We have presented evidence that representations of the changes between successive sounds, long known to characterize pitch perception, characterize perception of other dimensions as well. Contours derived from other dimensions seem qualitatively similar to those derived from pitch, as evidenced by the ease with which patterns can be matched across dimensions, and the extent to which these contours can be used to recognize melodies. Key aspects of relative pitch thus appear to be general features of the auditory system.

**Acknowledgments**—The authors thank Evelina Fedorenko, Nancy Kanwisher, Christophe Micheyl, Tali Sharot, and Nathan Witthoft for helpful comments on the manuscript. This work was supported by National Institutes of Health Grant R01 DC 05216.

## REFERENCES

- Attneave, F., & Olson, R.K. (1971). Pitch as a medium: A new approach to psychophysical scaling. *American Journal of Psychology*, *84*, 147–166.
- Dowling, W.J. (1978). Scale and contour: Two components of a theory of memory for melodies. *Psychological Review*, *85*, 341–354.
- Dowling, W.J., & Fujitani, D.S. (1970). Contour, interval, and pitch recognition in memory for melodies. *Journal of the Acoustical Society of America*, *49*, 524–531.
- Gagne, J., & Zurek, P.M. (1988). Resonance-frequency discrimination. *Journal of the Acoustical Society of America*, *83*, 2293–2299.
- Hyde, K.L., & Peretz, I. (2004). Brains that are out of tune but in time. *Psychological Science*, *15*, 356–360.
- Kochanski, G., Grabe, E., Coleman, J., & Rosner, B. (2005). Loudness predicts prominence: Fundamental frequency lends little. *Journal of the Acoustical Society of America*, *118*, 1038–1054.
- Liegeois Chauval, C., Peretz, I., Babai, M., Laguitton, V., & Chauvel, P. (1998). Contribution of different cortical areas in the temporal lobes to music processing. *Brain*, *121*, 1853–1867.
- MacMillan, N.A., Kaplan, H.L., & Creelman, C.D. (1977). The psychophysics of categorical perception. *Psychological Review*, *84*, 452–471.
- Marvin, E.W. (1995). A generalization of contour theory to diverse musical spaces: Analytical applications to the music of Dallapiccola and Stockhausen. In E.W. Marvin & R. Hermann (Eds.), *Concert music, rock, and jazz since 1945: Essays and analytic studies* (pp. 135–171). Rochester, NY: University of Rochester Press.
- McAdams, S., Winsberg, S., Donnadieu, S., De Soete, G., & Krimphoff, J. (1995). Perceptual scaling of synthesized musical timbres: Common dimensions, specificities, and latent subject classes. *Psychological Research*, *58*, 177–192.
- Melara, R.D., & Marks, L.E. (1990). Interaction among auditory dimensions: Timbre, pitch, and loudness. *Perception & Psychophysics*, *48*, 169–178.
- Micheyl, C., & Dai, H. (in press). A general area theorem for the same-different (2IAX) paradigm. *Perception & Psychophysics*.
- Moore, B.C.J., & Rosen, S.M. (1979). Tune recognition with reduced pitch and interval information. *Quarterly Journal of Experimental Psychology*, *31*, 229–240.
- Morris, R. (1993). New directions in the theory and analysis of musical contour. *Music Theory Spectrum*, *15*, 205–228.
- Neuhoff, J.G., Kramer, G., & Wayand, J. (2002). Pitch and loudness interact in auditory displays: Can the data get lost in the map? *Journal of Experimental Psychology: Applied*, *8*, 17–25.
- Neuhoff, J.G., McBeath, M.K., & Wanzie, W.C. (1999). Dynamic frequency change influences loudness perception: A central, analytic process. *Journal of Experimental Psychology: Human Perception and Performance*, *25*, 1050–1059.
- Patel, A.D. (2008). *Music, language, and the brain*. Oxford, England: Oxford University Press.
- Peretz, I., & Coltheart, M. (2003). Modularity of music processing. *Nature Neuroscience*, *6*, 688–691.
- Plantinga, J., & Trainor, L.J. (2005). Memory for melody: Infants use a relative pitch code. *Cognition*, *98*, 1–11.
- Semal, C., & Demany, L. (2006). Individual differences in the sensitivity to pitch direction. *Journal of the Acoustical Society of America*, *120*, 3907–3915.
- Stewart, L., von Kriegstein, K., Warren, J.D., & Griffiths, T.D. (2006). Music and the brain: Disorders of musical listening. *Brain*, *129*, 2533–2553.
- Trehub, S.E., Bull, D., & Thorpe, L.A. (1984). Infants' perception of melodies: The role of melodic contour. *Child Development*, *55*, 821–830.
- Verschuere, J., & van Meeteren, A.A. (1975). The effect of intensity on pitch. *Acustica*, *32*, 33–44.
- Viemeister, N.F., & Bacon, S.P. (1988). Intensity discrimination, increment detection, and magnitude estimation for 1-kHz tones. *Journal of the Acoustical Society of America*, *84*, 172–178.
- Warrier, C.M., & Zatorre, R.J. (2002). Influence of tonal context and timbral variation on perception of pitch. *Perception & Psychophysics*, *64*, 198–207.

(RECEIVED 1/24/08; REVISION ACCEPTED 5/9/08)