In press, *Psychological Science*

Is Relative Pitch Specific to Pitch?

Josh H. McDermott, Andriana J. Lehr, & Andrew J. Oxenham
Department of Psychology, University of Minnesota

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), and that contours in one dimension can be readily recognized in other dimensions, implicating similar or common representations. Most surprisingly, 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 this manipulation preserves the relationships between notes despite changing their individual pitches. Young infants share this ability to recognize melody transpositions (Plantinga & Trainor, 2005; Trehub, Bull, & Thorpe, 1984), suggesting that 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.

Despite its importance, little is known about how relative pitch is extracted. In music cognition it is common to distinguish between two components of relative pitch: contour and interval (Dowling & Fujitani, 1970) (Fig. 1a). For randomly generated novel melodies, recognition is dominated by the contour – the sequence of pitch change directions from note to note, containing the sign of a change but not its magnitude (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 at high levels, 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 sought to test whether comparable relative representations could be produced for other dimensions of sound. Our studies 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 apply to 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. Brightness varies across vowels and instruments, and can be varied orthogonally to pitch and loudness (Fig. 1b). Variation in both loudness and brightness are common in everyday stimuli. Rapid changes in spectral envelope characterize sequences of vowels, and local changes in intensity often signal stresses and accents in speech (Kochanski, Grabe, Coleman, & Rosner, 2005). Patterns in dimensions other than pitch are thus not without ecological validity, but it remains to be seen whether such variations can 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 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.

Method

We began our studies with the classic paradigm used by Dowling and colleagues (Dowling, 1978; Dowling & Fujitani, 1970) to demonstrate the importance of the melodic contour. They presented subjects with a randomly generated 5-note melody, followed by a second 5-note melody that was transposed up or down (Fig. 1c). The contour and interval composition of the two melodies could either be the same or different. They found that recognition depended primarily on the contour.

We extended the Dowling 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 of each condition the two stimuli had the same contour; on the other half they were different. When different, 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 based on whether the contour was in fact the same or different. These quantities were converted into an 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; 0.5 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, 2008), a task that would be less intuitive and slower to administer. Because we were interested in what subjects would naturally perceive, rather than 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 0.5.

Experiment 1: Contours in Dimensions Other Than Pitch

To test 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) (Fig. 1c). For each dimension the second stimulus could either be an exact transposition of the first, or could have the same contour but altered intervals (six conditions, 3x2 design). In this latter case the second melody had intervals that were the same sign as those from the first melody, but that always differed in magnitude. Subjects completed three blocks of 10 trials per condition, randomly ordered. 17 undergraduates (13 females; mean age of 21.82 years, SE = 1.24; averaging 6.2 years music training, SE = 1.34) participated.

Stimulus Details

The melodies were generated by choosing four intervals from the set (-2, -1, 1, 2 semitones) with equal probability, sampling 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 six semitones above 100 Hz (141.4 Hz). The spectral envelope was a Gaussian centered at 1000 Hz with 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, the spectral envelope was scaled in proportion to the center frequency (standard deviation was set to 25% of the centroid), normalizing to maintain constant rms amplitude. To partially compensate for differences in discrimination thresholds between dimensions, the shifts in spectral centroid were twice the size of the shifts in F0 in semitones. The lowest notes of the first and second brightness melodies always had spectral envelopes centered at 1000 Hz and 2119 Hz, respectively. The note level for the pitch- and brightness varying stimuli was 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 from 65 to 80 dB (limited to a 15 dB range to avoid uncomfortably loud levels). To maximize the discriminability of the contours, the range of the melody was mapped 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, presented back-to-back, with a 300-ms gap between melodies.

As can be seen in Figure 2a, subjects were substantially above chance at recognizing transposed patterns in all three dimensions – 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, F(1,16) = 11.32, p = .004, $h_p^2 = .41$, with no interaction, F(2,32) = 2.17, p = .131, 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. Subjects can thus recognize transpositions of loudness and brightness patterns nearly as well as transpositions of pitch patterns. The stimulus changes were not equated for discriminability, and because just-noticeable differences (JNDs) for shifts in spectral centroid and intensity are much larger, relative to their dynamic range, than for F0, any advantage for pitch could be due to this. The results display the hallmarks of a contour representation – invariance across transposition, and sensitivity more to the sign of change than the magnitude.

Experiment 2: Recognizing Pitch Contours in Other Dimensions

The apparent presence of contour-like representations for brightness and loudness raises the question of whether a pitch contour can be recognized when replicated in another dimension. To test this we ran an experiment in which the first stimulus was always defined by pitch variation, but where 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 for each dimension, one for each of the possible mappings. Because the scaling of a pattern ought not to matter much for contour extraction so long as the changes are readily perceived, we mapped each pitch melody to as large a range of the other dimension as was practical, preserving the relative size of the intervals.

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 (19 females; mean age of 22.63 years, SE = .82; averaging 4.93 years music training, SE = .83) participated.

As in Experiment 1, subjects performed best when the test melody was itself defined by pitch, $t(28) = 15.76$, $p<.0001$, and were little impaired when the intervals were altered, $t(28) = 13.0$, $p<.0001$, as shown in 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 and second melody could be defined by pitch, brightness, or loudness, yielding 9 conditions (3x3 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 (16 females; mean age of 22.15 years, SE = 1.2; averaging 5.55 years music training, SE = 1.01) participated.

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

One explanation for these results might be that subjects convert each contour to a verbal description (e.g. "up-down-down-up"), and then compare these descriptions, rather than comparing the contours themselves. The fast pace of the melodies and the short interval separating them made this a priori 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 20 subjects reported having used such a strategy, and then only for a minority of the time (less than half of the trials in all 5 cases). Verbal recoding thus seems unlikely to play a significant role.

Experiment 4: Familiar Melody Recognition

As a further test of the equivalence of contours in different dimensions, we tested recognition of 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 two, preserving the contour but altering the identity of its intervals. Interval stretching is less meaningful for brightness and loudness, as it is not obvious that there is any correspondence between interval sizes across dimensions; these conditions serve 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 level intervals [64 82] (unstretched) and [46 82] (stretched) dB SPL. 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 in random order and with a different assignment of conditions, such that each melody was played once per condition (there were three other conditions not discussed here). Thirty undergraduates (22 females; mean age of 21.83 years, SE = 0.8; averaging 3.3 years music training, SE = 0.71) participated.

Precise intervals are known to be important for familiar melody recognition, and our results confirm this, as stretching the intervals in the pitch stimuli impaired recognition – $t(29) = 6.03$, $p<.0001$ (paired t-test); see Figure 3. Performance in the stretched pitch condition was nonetheless better than with rhythmic cues alone, $t(29) = 8.97$, $p<.0001$ (paired t-test), indicating that contour alone can partially support recognition. Remarkably, performance in the non-pitch conditions was at similar levels. Post hoc comparisons (paired t tests, Bonferroni corrected) revealed the main effect of condition, $F(6,174) = 35.188$, $p<.0001$, $h_p^2 = .55$, to be driven by differences between the first (intact pitch intervals) and last (rhythm only) conditions and the other five conditions, none of which differed significantly from each other ($p > .05$). By comparison, there were no differences across condition in the proportion of incorrect responses (omitting non-responses), $F(6,174) = .96$, $p = .46$, ruling out criterion differences as accounting for 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 1-3, 4-6, and 7-9 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 6-9, SE = 2.05%), but no interaction with condition, $F(7.55, 218.87) = .49$, $p=.85$. Even when listeners knew little about what melodies to expect, loudness and brightness contours were nearly as useful as pitch contours for music recognition.

Pitch Matching Control Experiment

One trivial explanation of these results is that changes in spectral envelope and intensity affect a sound's pitch; under this hypothesis subjects would merely be recognizing a pitch contour produced by changes in these dimensions. To control for this 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 a button-click. Once an adjustment was made the cycle repeated, continuing until subjects indicated via button-click that the pitches matched. 20 such trials per condition were run, in random order. There were six conditions, four contrasting high and low levels (55 and 75 dB SPL, with complex tones or noise bursts as standards), and two contrasting high and low spectral envelopes (spectral centroids of 1000 Hz or 1682 Hz, with complex tone standards). The complex tones had missing F0s, drawn randomly from 50-100 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 in Experiment 1.

The effect on pitch was measured by the difference between the pure tone match frequency and the F0 of the complex tone, or the center frequency of the noise band. Fourteen undergraduates (9 females; mean age of 23.57 years, SE = 1.74; averaging 4.6 years music training, SE = 1.05) participated.

As shown in Figure 4, there was little effect of intensity on pitch matches – tones: $F(1,13) = 1.41$, $p=.26$; noise: $F(1,13) = .17$, $p = .69$. As can be seen from the histograms, pitch errors occurred occasionally for the tones, and were common for the noise bursts, but were no more frequent when the stimuli were high in level compared to when they were low. This 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, $F(1,13) = 6.95$, $p=.02$, $h_p^2 = .35$, again consistent with prior work (Warrier & Zatorre, 2002). Inspection of the histograms 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 conditions were within a semitone of the true F0. It seems unlikely that this weak effect could be responsible for the robust contour discrimination in Experiments 1-4, though it is difficult to completely rule out a small role. 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 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 loudness and brightness as well, implicating a representation of the changes between successive sounds rather than their absolute values. There were not obvious qualitative differences between dimensions. Indeed, contours in one dimension could be recognized when replicated in a different dimension, suggesting there may be a common representation for contours in different dimensions, or at least representations that are sufficiently similar to permit easy comparisons. This notion is underscored by our results with familiar melody recognition, for which the contour from brightness or loudness was almost as effective as a pitch contour.

It remains unclear whether this similarity between dimensions extends to the perception of intervals. Fully resolving this issue will require further studies specifically targeting interval perception in other dimensions. 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 whether two distinct but similar representations are involved. 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 pitch change discrimination (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 the dependence 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 decreases. This effect is reminiscent of prior reports of interference between pitch, brightness, and loudness (Melara & Marks, 1990; Neuhoff, Kramer, & Wayand, 2002). It remains to be seen whether this dependence is a function of learned associations (Neuhoff, McBeath, & Wanzie, 1999), for instance from covariation between these dimensions in speech and music, and/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? As discussed above, comparably precise interval information may not be available for other dimensions of sound, perhaps limiting 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 this range would be even smaller. As a result, pitch variation would appear 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, apply to 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.

Acknowledgements - The authors thank Evelina Fedorenko, Nancy Kanwisher, Christophe Micheyl, Tali Sharot, and Nathan Witthoft for helpful comments on the manuscript. The work was supported by NIH 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*(2), 147-166.

Dowling, W. J. (1978). Scale and contour: Two components of a theory of memory for melodies. *Psychological Review, 85*(4), 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*(2), 524-531.

Gagne, J., & Zurek, P. M. (1988). Resonance-frequency discrimination. *Journal of the Acoustical Society of America, 83*(6), 2293-2299.

Hyde, K. L., & Peretz, I. (2004). Brains that are out of tune but in time. *Psychological Science, 15*(5), 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*(2), 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*(10), 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. H. Marvin, R. (Ed.), *Concert Music, Rock, and Jazz since 1945: Essays and Analytic Studies*. Rochester: 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*(2), 169-178.

Micheyl, C., & Dai, H. (2008). A general area theorem for the same-different (2IAX) paradigm. *Perception & Psychophysics, In press*.

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*(1), 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*(4), 1050-1059.

Patel, A. D. (2008). *Music, Language, and the Brain*. Oxford: Oxford University Press.

Peretz, I., & Coltheart, M. (2003). Modularity of music processing. *Nature Neuroscience, 6*(7), 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*(6), 3907-3915.

Stewart, L., von Kriegstein, K., Warren, J. D., & Griffiths, T. D. (2006). Music and the brain: disorders of musical listening. *Brain, 129*(10), 2533-2553.

Trehub, S. E., Bull, D., & Thorpe, L. A. (1984). Infants' perception of melodies: The role of melodic contour. *Child Development, 55*(3), 821-830.

Verschuure, 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*(2), 198-207.

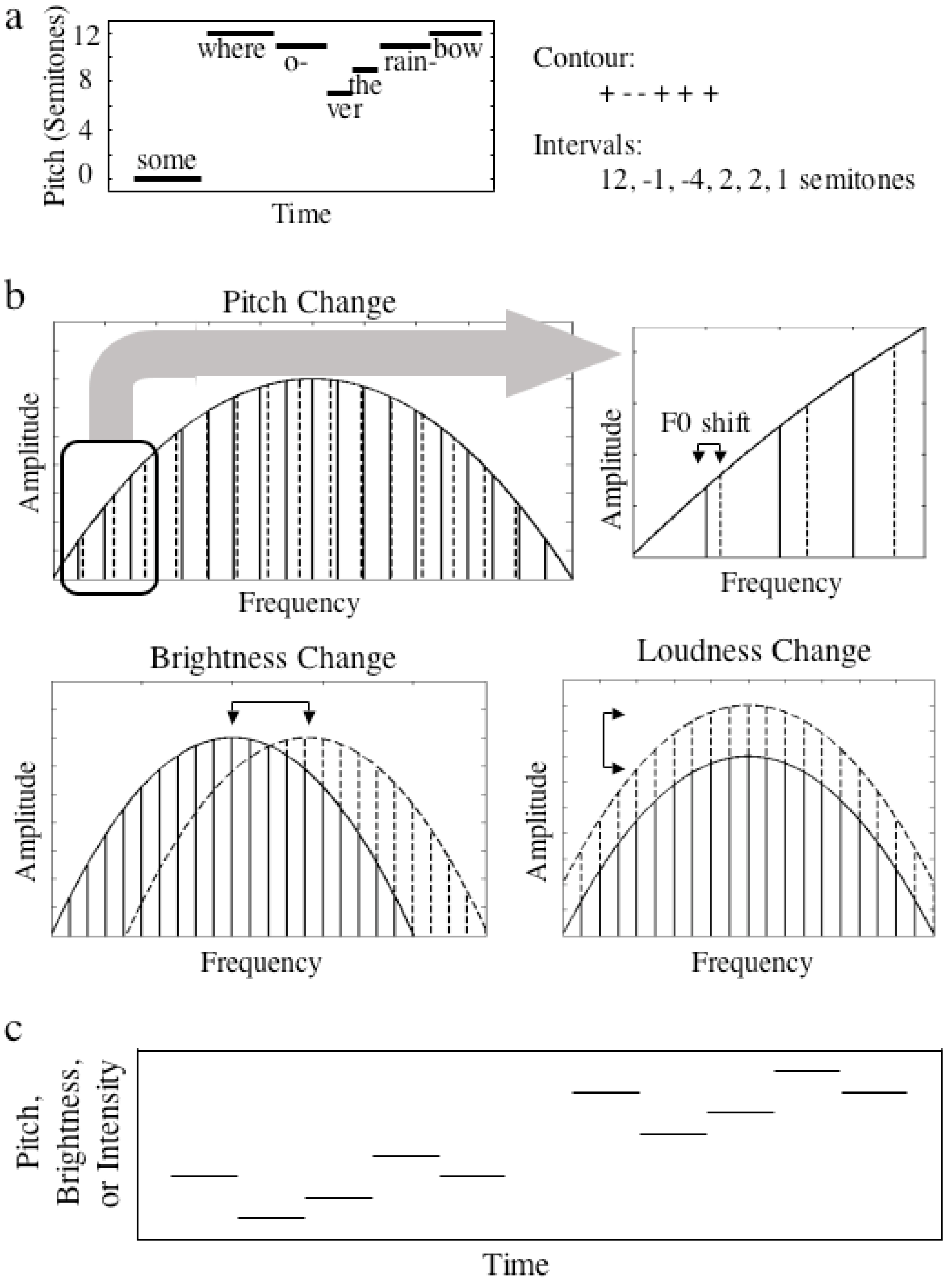

Fig. 1. Stimuli. (a) Diagram of the first few notes of "Somewhere Over the Rainbow", with contour and intervals specified. (b) Pitch, brightness and intensity can be varied independently. Straight lines denote frequency components of note; curved line is the spectral envelope. Solid and dashed lines denote two different sounds. Arrows indicate the stimulus change producing each perceptual change (F0 shift for pitch, spectral envelope shift for brightness, amplitude scaling for loudness). (c) Schematic of stimulus for a single trial of Experiment 1.

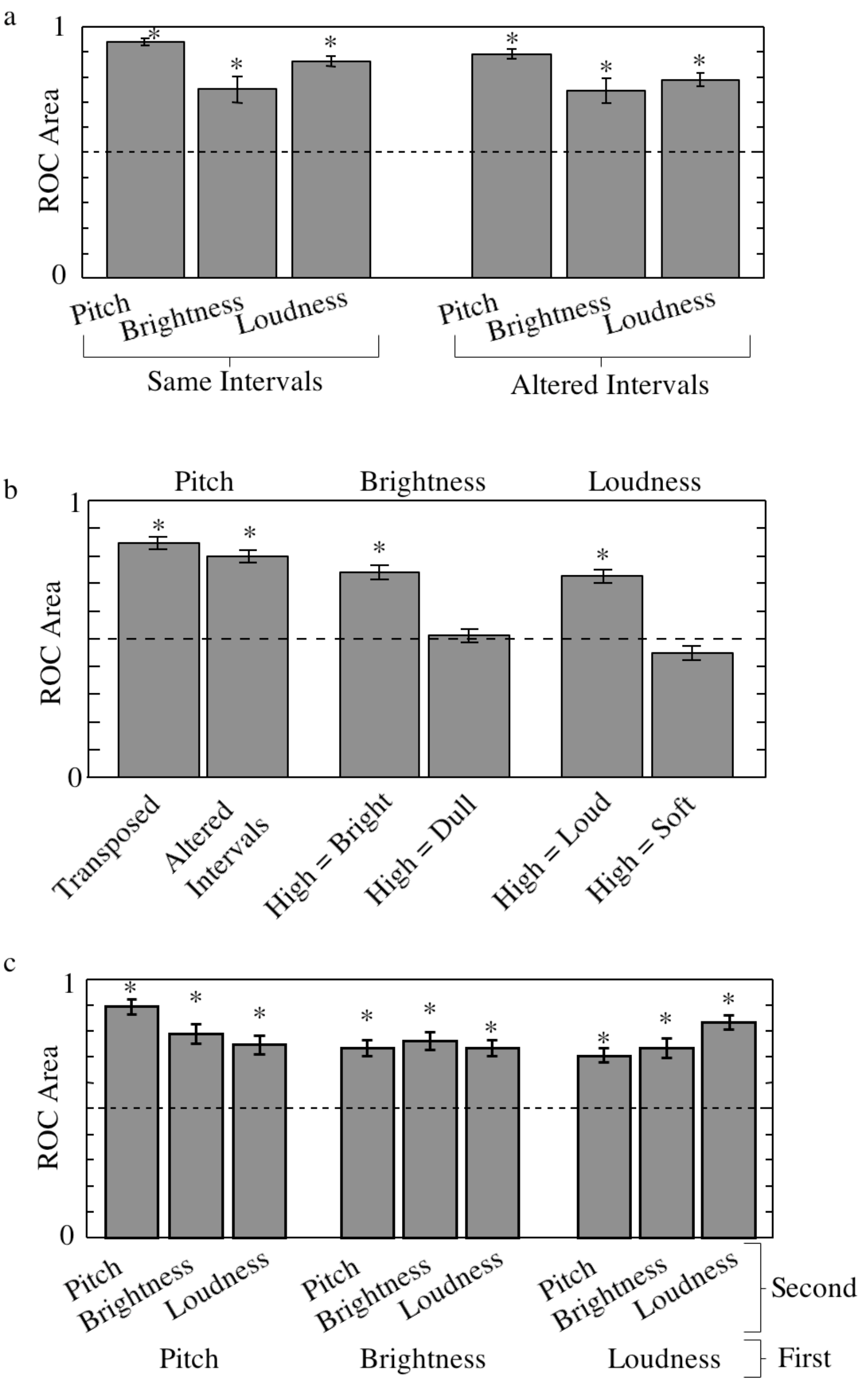

Fig. 2. Results for Experiments 1- 3. (a) Experiment 1 - Transpositions in brightness and loudness are recognized nearly as well as those in pitch. (b) Experiment 2 - Pitch contours are recognized when replicated in brightness and loudness. (c) Experiment 3 - Contours in one dimension can be recognized when replicated in another dimension. Dashed lines represent the level of chance performance. Here and elsewhere error bars denote SEMs; asterisks denote significance at the p<.001 level. In (c), the lower/upper labels below the graph denote the dimension of the first/second stimulus for a given condition.

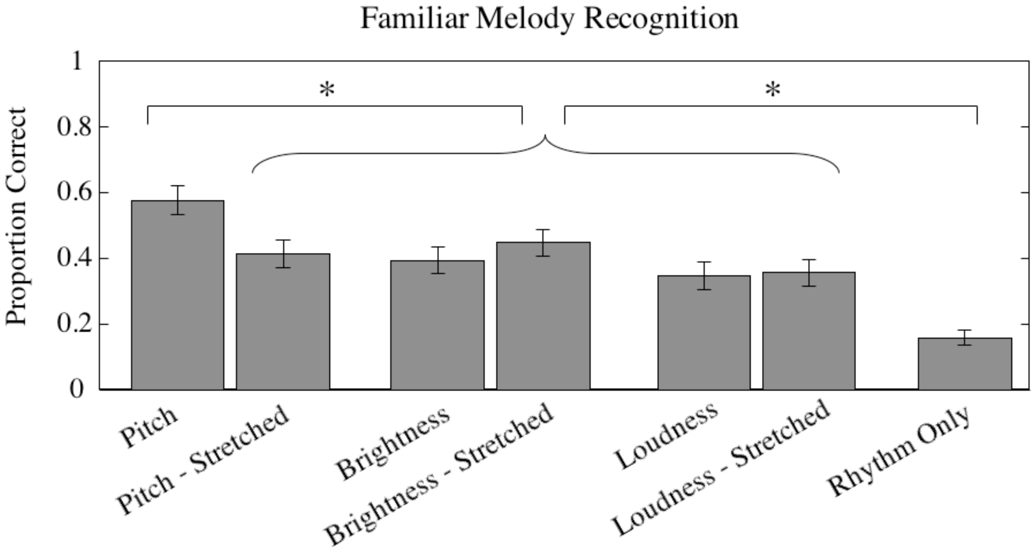

Fig. 3. Proportion of familiar melodies correctly identified when played in pitch, brightness, or loudness, with or without stretched intervals. The middle five conditions were statistically indistinguishable, indicating that contours from loudness and brightness were about as useful for recognition as pitch contours, once interval information was discounted.

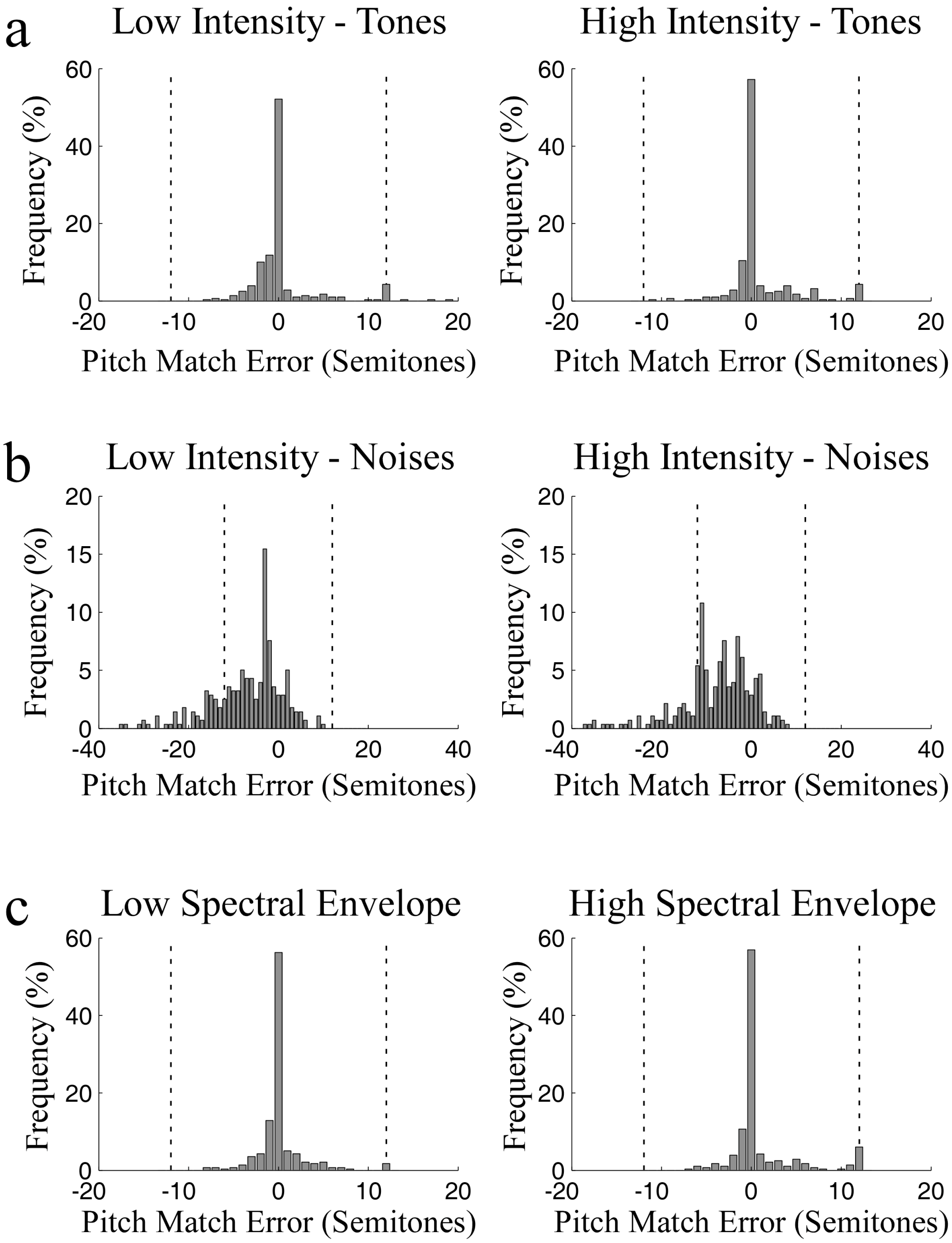

Fig. 4. Effect of intensity and spectral envelope on pitch, as measured with pure tone pitch matches to complex tones and noise bursts. (a) Matches to complex tones at low and high intensity. (b) Matches to noise bursts at low and high intensity. (c) Matches to complex tone standards with low and high spectral envelope centroids. Panels display the difference between the pure tone match frequency, and the F0 of the tone or the center frequency of the noise burst, measured in semitones. Histogram bins are one semitone wide, centered at integer numbers of semitones. Dashed lines denote one octave above and below the true F0 (or center of band for noise stimuli).