Journal of Experimental Psychology: Human Perception and Performance

Tipping the Scales: Auditory Cue Weighting Changes Over Development

Sarah C. Creel

Online First Publication, February 17, 2014. http://dx.doi.org/10.1037/a0036057

Journal of Experimental Psychology:
Human Perception and Performance

0096-1523/14/$12.00 DOI: 10.1037/a0036057

Tipping the Scales: Auditory Cue Weighting Changes Over Development

Sarah C. Creel
University of California San Diego

How does auditory processing change over development? This study assessed preschoolers' and adults' sensitivity to pitch contour, pitch height, and timbre in an association-memory paradigm, with both explicit (overt recognition) and implicit measures (visual fixations to melody-linked objects). In the first 2 experiments, child and adult participants associated each of 2 melodies with a cartoon picture, and recognition was tested. Experiment 1 pitted pitch contour cues against pitch height cues, and Experiment 2 pitted contour cues against timbre cues. Although adults were sensitive to multiple cues, children responded predominantly based on pitch height and timbre, with little sensitivity to pitch contour. In Experiment 3, however, children detected changes to all 3 cues well above chance levels. Results overall suggest that contour differences, although readily perceptible, are less memorable to children than to adults. Gradual perceptual learning over development may increase the memorability of pitch contour.

Keywords: pitch perception, auditory development, pitch contour, pitch height, timbre

How and why does perception change across development? In the auditory domain, we know that speech sound processing changes across developmental time with exposure to the native language (e.g., Werker & Tees, 1984). Less is known about other aspects of auditory perception. Particularly important for music perception is relative pitch—processing relations between successive pitches (pitch contours or pitch intervals; e.g., Bartlett & Dowling, 1980) rather than the absolute pitches themselves. Processing pitch relations is particularly important because these relations distinguish melodies from each other: Happy Birthday is still Happy Birthday whether it is sung by a soprano or a bass (a large difference in absolute pitch), and whether it is played on a kazoo or a xylophone (a large difference in timbre). How early in development is relative pitch a central factor in children's music representations?

At least three perspectives in the literature relate to this question. First is a *developmental stability* perspective: Children are sensitive to relative pitch, particularly contour, beginning in infancy (Plantinga & Trainor, 2005; Trehub, Bull, & Thorpe, 1984) and continuing into adulthood (Bartlett & Dowling, 1980). Second is *developmental shift*: Early in life, children are more sensitive to absolute pitch than to relative pitch, gradually shifting toward relative pitch processing (e.g., Saffran & Griepentrog, 2001). A third perspective is *developmental emergence*: Sensitivity to *numerous* musical properties, not just contour, emerges via experience with musical material (see, e.g., Creel & Jiménez, 2012 for a related account of voice recognition). Developmental emergence thus provides a *mechanism*—distributional learning—that can drive both overall changes in sensitivity during development, and developmental reweighting of cues (developmental shifts). The present study aimed to distinguish among these possibilities.

Correspondence concerning this article should be addressed to Sarah C. Creel, University of California San Diego, Department of Cognitive Science, 9500 Gilman Drive Mail Code 0515, La Jolla, CA 92093-0515. E-mail: creel@cogsci.ucsd.edu

Background

Despite a wealth of research on speech and music processing in early life, we know little about what the auditory experience of young children is like. Some studies suggest that a relative pitch focus is present in infancy (e.g., Plantinga & Trainor, 2005). Other accounts (Saffran & Griepentrog, 2001; Sergeant & Roche, 1973; Takeuchi & Hulse, 1993; see also Stalinski & Schellenberg, 2010) suggest a reprioritization during development, a shift away from absolute pitch processing toward relative pitch processing. Finally, other studies implicate other salient properties in the organization of musical representations, such as timbre (Trainor, Wu, & Tsang, 2004) and timing (Hannon & Trehub, 2005a, 2005b); however, few studies have investigated how these factors stack up against each other or against pitch cues.

Some studies suggest that relative pitch processing, at least at the level of pitch contour,[1] is available throughout development, along with mild sensitivity to absolute pitch (possibly implicit; Schellenberg & Trehub, 2003). Pioneering work by Trehub, Bull, and Thorpe (1984; Trehub, Thorpe, & Morrongiello, 1985) demonstrated that infants under a year old detect changes in the contour of a repeated melody at different pitch levels, suggesting sensitivity to relative pitch and, particularly, contour cues—the "ups" and "downs" of pitch direction in a melody (see also Plantinga & Trainor, 2005; Schellenberg & Trehub, 1996). On this view, mild sensitivity to absolute pitch information also remains present throughout life. This is supported by evidence that both children (Trehub, Schellenberg, & Nakata, 2008) and adults (Schellenberg & Trehub, 2003; Smith & Schmuckler, 2008) detect

[1] The current study discusses relative pitch in terms of pitch contour—the patterns of pitch rises and falls (increases and decreases in fundamental frequency) in a sequence of pitches. However, it is important to keep in mind that adult musical competence seems also to include pitch *interval* information, which is finer-grained than contour. Interval information keeps track of the exact pitch ratios between successive pitches rather than the binary up/down distinction. Thus the current study cannot distinguish between these two "grain sizes" of relative pitch information.

changes to the absolute pitch of familiar musical excerpts (see also Levitin, 1994, on adult pitch production). This work implies developmental stability in pitch processing, with a consistent focus on relative pitch.

Other researchers suggest a priority for absolute pitch information early in life. The basic thesis is that children start out perceiving pitch in absolute terms, and develop relative pitch later (Saffran & Griepentrog, 2001; Takeuchi & Hulse, 1993). This is supported by an apparent critical or sensitive period for developing absolute pitch labeling abilities: If one is not exposed to music lessons before the age of 9 or so, absolute pitch naming is unlikely to develop (Sergeant, 1969; Sergeant & Roche, 1973; Ward, 1999). Some empirical evidence also supports a processing shift. Saffran and Griepentrog showed that infants, but not adults, appeared to segment statistical tone streams using absolute pitch information, yet adults, but not infants, segmented streams using relative pitch information. Note, however, that Trainor (2005) offers a critique of Saffran and Griepentrog's finding that infants do not use relative pitch. Specifically, the test items those infants heard were identical in bigram (two notes in sequence) frequency to the training in both conditions. Thus, if children were attending to the relative pitch of bigrams only (rather than trigrams), they would find both "legal" and "illegal" test items equally acceptable (Trainor, 2005). Thus, Trainor argues, this does not constitute evidence of poor relative pitch in infants. Stalinski and Schellenberg (2010) presented children and adults with pairs of melodies and asked them to rate the similarity of each pair. They showed that 5- to 7-year-old children judged transpositions of the standard melody—a change in absolute pitch but matched in contour—to be as dissimilar as two completely different melodies, unlike older children or adults. Of course, children could have been responding in the "absolute pitch" trials based on absolute pitch, or based on average relative pitch (the notes in the second melody are on average higher in pitch than the notes in the first melody; see Creel & Tumlin, 2012). Taken together, these studies hint at a change in auditory encoding away from an absolute pitch (or average-relative-pitch) reference frame, toward a relative-pitch-based reference frame.

A third area of research suggests that children may process music, like other patterned information, differently than adults simply because they have had less auditory experience. An adult-like amount of auditory/musical experience is needed to both sharpen and reweight a variety of musical cues—absolute pitch, relative pitch, timbre, timing—toward adult levels. For instance, Trehub, Schellenberg, and Nakata (2008) found evidence that absolute pitch encoding may *improve* over development—not only contradicting a developmental decline in absolute pitch processing, but suggesting experiential tuning of absolute pitch representations. Although infant studies suggest sensitivity to contour (Plantinga & Trainor, 2005; Trehub et al., 1984), experiments with older children (preschool-aged and up) in more taxing paradigms find that children are relatively poor at utilizing pitch contours in processing emotional speech (Quam & Swingley, 2012; see related work by Morton & Trehub, 2001; Nelson & Russell, 2011), distinguishing words (Quam & Swingley, 2010), and recognizing newly learned voices (Creel & Jiménez, 2012), suggesting that there are slow developmental increases in pitch processing with increasing experience. These studies are also consistent with developmental sharpening in other types of auditory processing: Preschoolers are not yet adult-like at processing cues to speech sounds (e.g., Ohde & Haley, 1997), or at identifying voices from spectral information (Creel & Jiménez, 2012; Mann, Diamond, & Carey, 1979). These studies together imply that children require lengthy perceptual learning to encode numerous auditory cues including pitch contour, or to map pitch contour information to representations (emotional states, individuals). This suggests there may be lengthy improvement in music perception, and pitch contour in particular, over development.

Of course, conflicting findings of contour sensitivity in infants (e.g., Trehub et al., 1984) but weak pitch contour processing in young children (e.g., Quam & Swingley, 2010, 2012) are difficult to assess as they typically use different paradigms. Infant paradigms, such as habituation and conditioned head-turn change detection, which both rely on short-term memory (STM) representations, suggest early (infant) sensitivity to pitch contour. However, accuracy and similarity judgment data (e.g., Stalinski & Schellenberg, 2010) suggest that children are less sensitive than adults to pitch relations. Does this mean that contour sensitivity has a U-shaped function (good in infancy and adulthood, poor in between), or are these seeming developmental changes an artifact of differences in the paradigms used? Further, although a few studies have pitted absolute pitch cues against relative pitch (contour) cues, almost none have examined how pitch contour—the basis of a melody's identity—fares against other auditory features. Thus, it is unknown what the salient dimensions are in children's musical representations.

Music and Eye Tracking

One technique that is relatively versatile across age groups is the "visual world" eye tracking paradigm (Cooper, 1974; Tanenhaus, Spivey-Knowlton, Eberhard, & Sedivy, 1995). Although it is typically used to study language in adults (Tanenhaus et al., 1995) and children (Trueswell, Sekerina, Hill, & Logrip, 1999), Creel and Tumlin (2012) recently adapted the visual world paradigm to study music processing. They trained adult listeners to associate each of several brief melodies with a particular nonsense shape, like musical "words" for the shapes. Creel and Tumlin then used visual fixations to the shapes to measure recognition moment-by-moment as the melody unfolded in time. Some pairs of melodies had identical absolute and relative pitch up to the final note (absolute pitches: $C_4D_4E_4F_4G_4$ and $C_4D_4E_4F_4E_4$; relative pitches, ↑M2 ↑M2 ↑m2 ↑M2 and ↑M2 ↑M2 ↑m2 ↓m2), yet others matched only in relative pitch up to the final note (absolute pitches: $G_4E_4F_4D_4C_4$ and $C^{\#}_5A^{\#}_4B_4G^{\#}_4A^{\#}_4$; relative pitches: ↓m3 ↑m2 ↓m3 ↓M2 and ↓m3 ↑m2 ↓m3 ↑M2; numbers denote octaves where 4 = middle C upward, 5 = C above middle C upward, and so forth). After training to 90% accuracy, adults were tested: They saw two shapes at a time, and heard the melody that matched one of them. Eye movements to both pictures were measured as the melody played. If listeners were using absolute pitch to distinguish melodies, then the ones differing in absolute pitch should be differentiable from each other earlier than those matching in absolute pitch, generating sooner looks to the correct picture. If listeners used only relative pitch, they should not show recognition until the last note in the melody (i.e., a different interval/pitch direction). Consistent with use of absolute pitch information, adults looked toward the correct picture sooner when the two pictures' melodies differed in absolute pitch ($G_4E_4F_4D_4C_4$ and $C^{\#}_5A^{\#}_4B_4G^{\#}_4A^{\#}_4$) than when they did not.

Further experiments in Creel and Tumlin (2012) teased apart the factors contributing to the results in their first experiment. Those experiments verified that adults are still sensitive to absolute pitch (as in Schellenberg & Trehub, 2003), as well as what Creel and Tumlin termed global relative pitch (knowing that a melody is high-pitched *relative to* the pitches heard in the experiment—on analogy to Navon's 1977 visual stimuli, e.g., large S's composed of small H's). In any case, Creel and Tumlin's study suggests that an eye-tracked melody-object mapping paradigm may in principle be a fruitful way to explore melody processing at a range of ages, as it does not require a verbal or manual response.

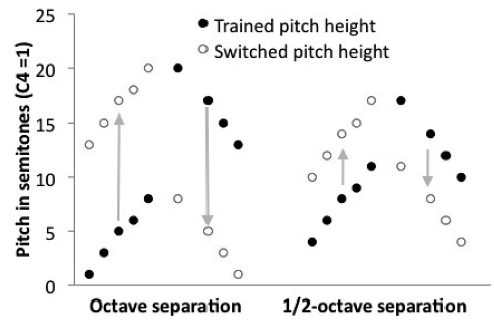


Figure 1. Examples of melodies used in Experiment 1. Solid points indicate the melodies participants were trained on. Hollow points indicate novel melodies (changes in absolute pitch) that participants heard during Test Blocks 2 and 3.

The Current Study

The research reviewed above leaves open several questions. First, how do children (vs. adults) weight various auditory cues in representing melodies: relative pitch (implying developmental stability), pitch height or timbre (implying developmental shift or developmental emergence), or some combination of cues? Assessing the relative roles of various cues to musical identity is crucial for understanding how and whether auditory processing changes across development. Second, do children *lose* sensitivity to pitch contour after infancy, only to regain it in adulthood, or is pitch contour particularly difficult to encode in long term memory?

To explore these questions, children and adult controls were tested in a child-friendly paradigm previously used to study word learning (Creel, 2014) and voice learning (Creel & Jiménez, 2012), but using musical stimuli as in Creel and Tumlin (2012). Children were told that each of two cartoon creatures had a favorite song, and that they would see each creature and hear the creature's favorite song. After several repetitions of each favorite song, children saw both creatures at the same time, heard a song, and were asked to select the creature whose favorite song it was. This paradigm allowed collection of explicit responses (pointing to creatures) as well as nonverbal responses—visual fixations to melody-associated pictures. Experiment 1 juxtaposed relative pitch cues (melodic contour, the pitch rises and falls from note to note) and absolute pitch (pitch height) cues to melody identity. Experiment 2 juxtaposed instrumental timbre cues versus pitch contour cues. Finally, Experiment 3 directly examined children's ability to *discriminate* changes in contour, pitch height, and timbre.

Experiment 1

The first experiment tested children's ability to associate melodies differing in both relative and absolute pitch (pitch height) information with cartoon pictures. Pilot work showed that the subtly different melodies that adults in Creel and Tumlin (2012) learned readily (such as $C_4D_4E_4F_4G_4$ vs. $C_4D_4E_4F_4E_4$) were very difficult for children to learn. Therefore, the two melodies used here (see Figure 1) employed very different contours (one rising, one falling) and an additional duration/note number cue (the first note of the falling melody was twice the duration of the other notes). These two melodies, one with four rising pitch intervals and one with three falling pitch intervals, constitute a large contour difference—an edit distance of approximately four. This is an even larger change than was reliably detected by Trehub et al.'s (1985) 6- to 8-month-old infants in a conditioned head-turn paradigm (a change of one note in a six-note melody, which altered at most two note-to-note pitch contours). Further, melodies had a large difference in absolute pitch: the pitches of one melody were either one octave higher (a doubling of fundamental frequency) or one tritone higher (half an octave) than the other melody. Absolute pitch cues will be referred to in this experiment as *pitch height* as a more neutral term, so as not to presuppose that listeners are processing pitch absolutely versus in terms of global pitch relations.

Both child and adult listeners learned melodies in a series of 16 training trials (see Table 1). Next, in the first set of test trials, children were asked to point to the creature whose song they heard. Refresher training trials followed (eight), then more test trials (eight), then a third training-test cycle (8 + 8). The reason for repeating training trials was to assess whether learners might show accuracy increases with additional exposure. In the first round of test trials, only the melodies from the learning phase were played. In the additional later test trials (but not additional learning trials), listeners heard the original melodies plus new versions of the melodies with pitch height cues switched: the higher melody was heard in the lower pitch range, and the lower melody was heard in the higher pitch range.

If listeners identify melodies based on relative pitch contour, despite a change in pitch height, they should point and look to the correct creatures on switched-pitch-height ("switched") trials. However, if they have difficulty identifying melodies after a change in pitch height, accuracy should be lower on switched-pitch trials than on original-pitch trials.

Method

Participants. N = 40 preschool-aged children (*M* = 4.2 years, *SD* = 0.7, range: 2.9–5.5; 19 female) and 32 college-aged adults took part. The child samples throughout were obtained by contacting local preschools to assess interest level. Samples were approximately balanced for gender, and were predominantly Caucasian with other ethnicities represented. Socioeconomic status information was not obtained. Adults throughout were recruited via the university's human participant pool, and were predominantly Caucasian and Asian. One child participant was not included in eye tracking analyses due to computer error. One more

Table 1
Trial Sequences Used in all Experiments

Trial type	# Trials	Original stimuli	Altered stimuli
Training 1a	8	8	
Two distractor trials			
Training 1b	8	8	
Test Block 1	8	8	
Training 2	8	8	
Two distractor trials			
Test Block 2	8	4	4
Training 3	8	8	
Two distractor trials			
Test Block 3	8	4	4

child and one adult were excluded from eye tracking analyses due to data criteria described below.

Stimuli. Auditory stimuli were created in Finale, 2009.r2 software (MakeMusic, Inc.) and exported as .aiff files, which were then converted to .wav files and scaled to 70 dB amplitude in Praat 5.1.44 software (Boersma & Weenink, 2010). The two melodies were each recorded at four different pitch levels (see Figure 1), two for the octave separation (melody range: C_4–G_4, and C_5–G_5) and two for the half-octave separation ($E^b{}_4$–$B^b{}_4$ and A_4–E_5). Melodies were highly distinct in relative pitch (one rising; one falling) because pilot work suggested that young children had difficulty learning mappings with more similar melodies. All melodies were presented in the MIDI "voice oohs" timbre. A voice-like timbre was used in order to make melodies interesting to children (see Weiss, Trehub, & Schellenberg, 2012). Presentation rate was 400 ms per quarter note (total duration: 1,200 ms plus time for reverberation to fade). Visual stimuli were two cartoon creatures designed to be engaging to children and highly discriminable, previously used in Creel (2014) and Creel and Jiménez (2012).

Procedure. The experiment was run in a quiet area in the child's day care or preschool, or in a quiet room in the lab (adults). Each child sat in front of an experimental display monitor linked to a Mac Mini running OS 10.4.1 and Matlab (2008a). Directly beneath the monitor sat an Eyelink 1000 eye tracker (Mississauga, ON, CA; www.sr-research.com) running in remote mode, linked to a Dell tower. For children, the monitor for the Dell tower faced directly away from the experimental monitor to avoid introducing side biases into children's looking patterns.

Each participant first completed the eye tracker calibration sequence from the Eyelink Toolbox (Cornelissen, Peters, & Palmer, 2002), framed for children as a "follow-the-dot game." The experiment was then presented using Psychtoolbox3 for Matlab (Brainard, 1997; Pelli, 1997). Instruction screens appeared, and the experimenter read them aloud to the child. Adults read instructions silently.

In the learning phase, children were told (and adults read) that they would see two creatures, each of whom had a favorite song. On 16 training trials (eight per melody; see Table 1), participants saw a creature move onto the screen and pause at the center, at which point the creature's favorite song played over appropriately sized headphones (KidzGear headphones for children, Sennheiser Pro HD280 headphones for adults). The creature then moved offscreen. Next came eight test trials (Test Block 1). On each test trial, the two creatures appeared, static, on either side of a blank white screen. One melody played, and the child was asked to point to the creature whose favorite melody it was. Accuracy and visual fixations were measured on all test trials. The experimenter recorded the child's pointing direction with a mouse click. Adults indicated responses by clicking the mouse themselves.

Following the first test, there were another eight training trials, eight testing trials, eight more training, and eight more testing. Each set of refresher training trials was followed by two distractor trials (pictures of animals moving up and down with clapping and cheering noises) in order to remove any working memory traces of trained melodies prior to testing. Test Blocks 2 and 3 contained 50% original-pitch trials, and 50% switched-pitch trials where each melody was played in the absolute-pitch range of the other. This design allowed (a) assessment of learning prior to introducing switched-pitch trials, which might be potentially confusing; and (b) detection of improvement, or fatigue, with increased learning trials. For the most part, neither decreases nor increases in performance on learned stimuli (from Test Block 1 to Test Blocks 2–3) were observed. For maximum comparability, though, accuracy and visual fixations were compared between original and switched trials on Blocks 2–3 only, where both stimulus types were presented. Equal numbers of participants learned: low rising and high falling, versus high rising and low falling; octave separation versus half-octave separation.

Results

Accuracy. Throughout the study, correct responding was defined based on pitch contour. In the current experiment, if participants responded based on pitch height, answers in the switched-pitch condition would be consistently counted as *incorrect*. Adults performed with near perfect accuracy on original-pitch trials. Their responses differed on switched-pitch trials depending on the pitch height difference (Figure 2a, solid lines): When melodies were one octave apart, adults showed mixed responding, but when melodies were a half-octave apart, adults responded completely based on pitch contour. Children learned only modestly, and appeared to respond based on pitch height (Figure 2a, dashed lines). Children's accuracy in Block 1 was overall comparable to accuracy in Blocks 2 and 3 on original-pitch trials, suggesting they did not benefit from additional exposure. Because no learning effects were evident, analyses were restricted to Blocks 2 and 3 for maximum comparability.

Responses were empirical-logit (e-logit) transformed to correct for the non-normal distribution of accuracy data.[2] Transformed responses were subjected to an analysis of variance (ANOVA) with Block (2, 3) and Pitch Height (original, switched) as within-participants factors, and Pitch Separation (octave, 1/2 octave) as a between-participants factor. Each Age Group was examined separately.

Adults showed an effect of Pitch Height, with higher accuracy on original- than switched-pitch trials, $F(1, 30) = 19.91$, $p = .0001$. There was also an effect of Pitch Separation, $F(1, 30) = 7.77$, $p = .009$, with higher accuracy when melodies were separated by a smaller pitch distance (1/2 octave). Finally, there was an interaction of Pitch Height × Pitch Separation, $F(1, 30) = 17.20$, $p = .0003$. This interaction resulted from a significant effect of Pitch Height at the octave pitch separation, $F(1, 15) = 19.38$, $p =$

[2] Note that all reported statistics (both accuracy and visual fixations) held when analyses were run on raw data.

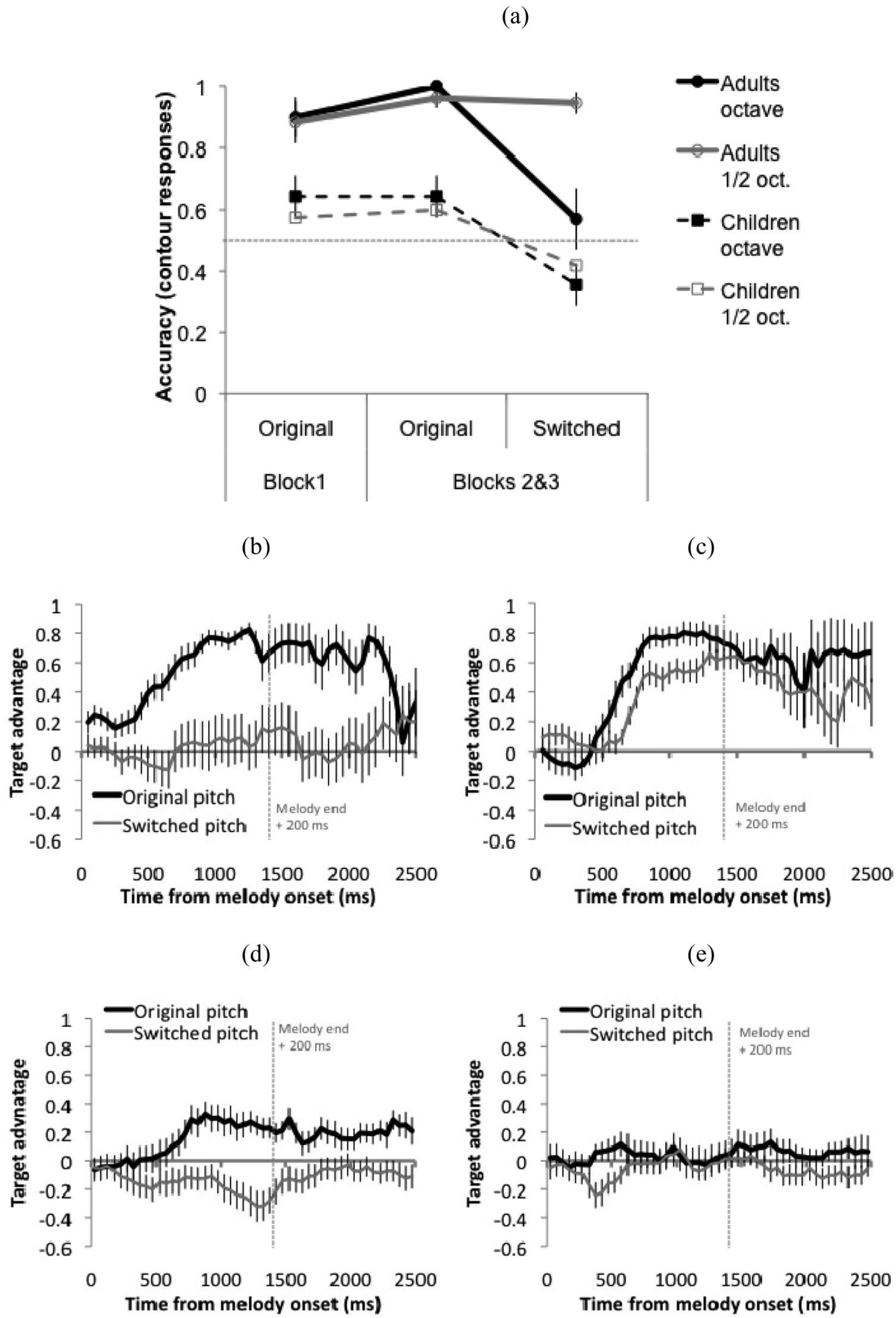


Figure 2. Experiment 1, (a) accuracy, and (b–e) looks to pictures in Blocks 2–3 for (b) adults, octave separation; (c) adults, half-octave separation; (d) children, octave separation; (e) children, half-octave separation. Error bars are standard errors.

.0005: Adults performed at ceiling in the original-pitch condition (1.00 ± 0.0), but at chance (.570 ± .400; $t(15) = 0.92$, $p = .37$) in the switched-pitch condition (individual response patterns are depicted in the Appendix). There was no effect of Pitch Height at the half-octave pitch separation ($F(1, 15) = 1.94$, $p = .18$; above chance overall, .953 ± .130, $t(15) = 13.75$, $p < .0001$). No other effects approached significance.

Children showed a different pattern: The only significant effect was Pitch Height, $F(1, 38) = 14.66$, $p = .0005$. This reflected above-chance accuracy on original-pitch trials (.622 ± .251;

$t(39) = 3.01$, $p = .005$) but *below*-chance accuracy—that is, responses based on pitch height rather than contour—on switched-pitch trials (.388 ± .213; $t(39) = 3.31$, $p = .002$). In fact, there was no evidence that children used pitch contour at all; if responses were rescored as correct based on pitch height, there was no *loss* of accuracy when the pitch contour information mismatched training (.622 vs. .613; $t(39) = 0.31$, $p = .76$). Children's age did not significantly predict their performance (all $p > .14$). This suggests that although adults' responses to conflicting pitch height and contour cues reflected the influence of both cues, children responded based on pitch height alone.

Visual fixations. Throughout the article, trials with fewer than 50% looks to the two pictures were dropped from analysis (in the current experiment, 17.9% of child trials and 13.3% of adult trials). This included looking to empty screen areas, looks offscreen, and times when the eye tracker could not find the child's eye. This last case occurred frequently for children as they began pointing, because their raised arms obscured the camera. Thus, the analysis window was confined to 0–1,200 ms after melody onset—the duration of the melody. On the assumption that humans take roughly 200 ms to plan and execute an eye movement based on external input (Hallett, 1986), this 0–1,200 ms window was shifted forward to 200–1,400 ms.

To assess real-time recognition of melodies, eye movements (Figure 2b–e) to the melody-matched picture (the "target") and the other picture were assessed. For melodies at each pitch level (original or switched), transformed looks to the other picture were subtracted from transformed looks to the target, creating a *target advantage* score. When target advantage is positive, there are more looks to the target picture than to the other picture. When it is negative, there are more looks to the other picture than to the target. Following Barr (2008), looks were empirical-logit transformed before computing target advantage.

A mixed ANOVA was performed on target advantage from 200–1,400 ms in Test Blocks 2 and 3, with Pitch Height as a within-participants factor and Pitch Separation as a between-participants factor. Data were collapsed over Block to yield greater power. As with accuracy analyses, each age group was analyzed separately.

Adults (Figure 2b–c) showed an effect of Pitch Height, $F(1, 29) = 39.52$, $p < .0001$, with greater target advantage on the original-pitch trials than switched-pitch trials. There was also a Pitch Height by Pitch Separation interaction, $F(1, 29) = 12.37$, $p = .001$. This resulted from different magnitudes of the Pitch Height effect at each level of Pitch Separation. For the octave pitch separation, Pitch Height was significant, $F(1, 14) = 31.73$, $p < .0001$, reflecting above-chance target advantage for original-pitch trials, $t(14) = 8.51$, $p < .0001$ but not for switched-pitch trials, $t(14) = 0.10$, $p = .93$. For the half-octave pitch separation, Pitch Height was also significant, $F(1, 15) = 8.03$, $p = .01$, with greater target advantage on original-pitch trials (above chance, $t(15) = 8.28$, $p < .0001$) and lower, but still above-chance, target advantage for switched-pitch trials, $t(15) = 7.57$, $p < .0001$.

Children (Figure 2d–e) showed a somewhat different pattern of visual fixations. There was an effect of Pitch Height, $F(1, 36) = 11.77$, $p = .002$, indicating overall higher target advantage on original-pitch trials. There was also a Pitch Height × Pitch Separation interaction, $F(1, 36) = 4.15$, $p < .05$. This interaction resulted from differing effects of Pitch Height at each level of Pitch Separation. For the larger (octave) pitch separation, Pitch Height was significant, $F(1, 18) = 9.21$, $p = .007$, reflecting above-chance target advantage on original-pitch trials, $t(18) = 2.47$, $p = .02$, but *below*-chance target advantage on switched-pitch trials, $t(18) = 2.63$, $p = .02$. For the smaller (half-octave) pitch separation, Pitch Height did not reach significance, $F(1, 18) = 2.58$, $p = .13$ and target advantage did not exceed chance overall, $t(18) = 0.29$, $p = .77$. Interestingly, the only significant correlation with child age was with switched-pitch target advantage, $r(38) = -.32$, $p = .04$. This negative correlation suggests that pitch height-based responding *increased* with age, counter to the idea of a developmental decrease. Thus, eye movements suggested that children, to the extent that they made systematic eye movements, did so based on pitch height rather than pitch contour.

Discussion

Adults showed sensitivity both to pitch contour and pitch height: Responses were swayed toward the pitch-height match when there was a large pitch separation (one octave), but not a smaller pitch separation (1/2 octave). Further, even for the more subtle pitch separation, eye movements to the target picture were slowed by a change in pitch height. Children, on the other hand, showed a singular dependence on pitch height: When pitch height changed, responses and looks reversed completely, with no hint of an effect of pitch contour. This pattern of results may mean that melodic contour is less salient to children than other types of pitch information, such as pitch height.

Also worth noting is that children are using *both* types of pitch cues—melodic contour, pitch height—somewhat weakly, judging by the modest accuracy levels in the current experiment. This low accuracy raises some questions about what children are doing in the task. Previous studies on word learning (Creel, 2014) and talker learning (Creel & Jiménez, 2012) suggest that preschoolers find this task quite easy when given highly discriminable stimuli (two distinct words, a male voice vs. a female voice). This implies that, rather than finding the task difficult, children may have difficulty forming distinct long-term representations of the two melodies. What, then, are they representing? Are other cues more recognizable, or more prominent in memory, than pitch contour? The next experiment explored the relative salience of pitch contour versus another cue: instrumental timbre. Although infants (Trainor et al., 2004) and adults (Creel, 2012; Halpern & Müllensiefen, 2008) are sensitive to timbre, its relative strength as a recognition cue has, to my knowledge, never been tested in children.

Experiment 2

In this experiment, participants were taught to recognize two melodies that differed in timbre. Melodies were much like those used in Experiment 1, but instead of a pitch height difference between rising and falling melodies, there was a timbre difference (see Figure 3). For half of participants, one melody was played in a muted-trumpet timbre, and the other a vibraphone timbre. For the other half, one melody was played in a bassoon timbre, and the other in an alto saxophone timbre. These instrumental timbres were selected to be, respectively, highly distinct or highly similar (see Figure 4) based on Iverson and Krumhansl's (1993) multidimensional scaling analysis of 16 timbres with adult listeners. On

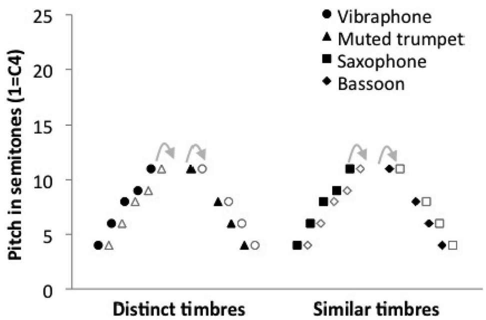


Figure 3. Examples of melodies used in Experiments 2. Point shape corresponds to timbre. Solid points indicate the melodies participants were trained on. Hollow points indicate novel melodies (changes in absolute pitch or in timbre) that participants heard during Test Blocks 2 and 3. Arrows indicate changes in timbre.

test trials, rather than switching pitch height, *timbres* were sometimes switched between melodies—that is, each melody was sometimes heard in the timbre of the other melody. If timbre is a relatively weak cue, then listeners should now respond based on pitch contour, regardless of timbre. However, if timbre is a stronger cue than pitch contour, then listeners should reverse their responses when the timbre is switched.

Method

Participants. N = 32 new preschool-aged children ($M = 4.6$ years, $SD = 0.5$, range: 3.7–5.6; 13 female) and 32 new college-aged adults from the same populations as in previous experiments took part. Two adult participants were dropped from eye tracking analyses for having less than two observations per cell after the trial exclusion criteria were applied.

Stimuli. The stimuli were the melodies used in Experiment 1, except that there were no differences between melodies in pitch height, just in timbre. Half of participants learned at the lower pitch level (E^b_4–B^b_4), half learned at the higher pitch level (A_4–E_5). Each melody was equally likely to be heard in either timbre, and each timbre pair (distinct pair, similar pair) was heard by an equal number of participants in each age group.

Procedure. This matched Experiment 1, except that the "switched" trials in Test Blocks 2 and 3 switched timbre rather than pitch height.

Results

Accuracy. Adults (Figure 5a, solid lines) approached ceiling accuracy on original-timbre trials. For the distinct-timbre condition, adults showed a numerical tendency toward timbre responses rather than contour responses; however, for the similar timbres they showed strong contour responses. Children (Figure 5a, dashed lines) responded strongly based on timbre in the distinct-timbre condition, with somewhat weaker responding in the similar-timbre condition. As in Experiment 1, ANOVAs were conducted on accuracy, with Block and Timbre (original, switched) as within-participants factors, and Timbre Separation as a between-participants factor.

Adults showed an effect of Timbre, $F(1, 30) = 31.81$, $p < .0001$, with higher accuracy on original-timbre trials than switched-timbre trials. An effect of Timbre Separation, $F(1, 30) = 27.47$, $p < .0001$ reflected higher accuracy for the similar-timbre condition than for the distinct-timbre condition. Finally, Timbre and Timbre Separation interacted, $F(1, 30) = 28.38$, $p < .0001$. The interaction resulted from a null effect of Timbre for similar timbres ($F(1, 15) = 0.70$, $p = .42$; above chance overall, $.965 \pm .079$, $t(15) = 20.53$, $p < .0001$), but a significant effect of Timbre for the distinct timbres, $F(1, 15) = 31.67$, $p < .0001$. Within the distinct-timbre condition, original-timbre trials exceeded chance accuracy ($.969 \pm .085$; $t(15) = 16.21$, $p < .0001$) although switched-timbre trials did not ($.375 \pm .411$; $t(15) = 1.25$, $p = .23$).

For children, a slightly different pattern emerged. There was an effect of Timbre, $F(1, 30) = 22.58$, $p < .0001$, with higher accuracy for original-timbre than switched-timbre trials. An effect of Timbre Separation, $F(1, 30) = 10.49$, $p = .003$ reflected higher accuracy in the similar-timbre condition than the distinct-timbre condition. Finally, the two factors interacted, $F(1, 30) = 12.04$, $p = .002$, a result of a null effect of Timbre for similar-timbre trials ($F(1, 15) = 2.56$, $p = .13$; above chance overall, $.629 \pm .219$, $t(15) = 3.04$, $p = .008$) but a significant effect of Timbre for distinct-timbre trials, $F(1, 15) = 21.24$, $p = .0003$. For distinct-timbre trials, children exceeded chance accuracy on original-timbre trials ($.758 \pm .290$; $t(15) = 3.66$, $p = .002$) but were

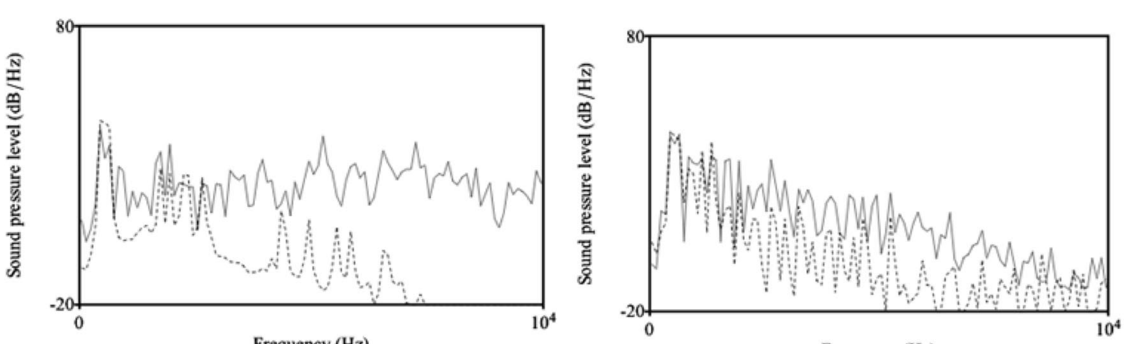


Figure 4. Long-term average spectrum of instrumental timbres used in Experiment 2, created in Praat (Boersma & Weenink, 2010). L: Solid = muted trumpet, dashed = vibraphone. R: Solid = alto saxophone, dashed = bassoon.

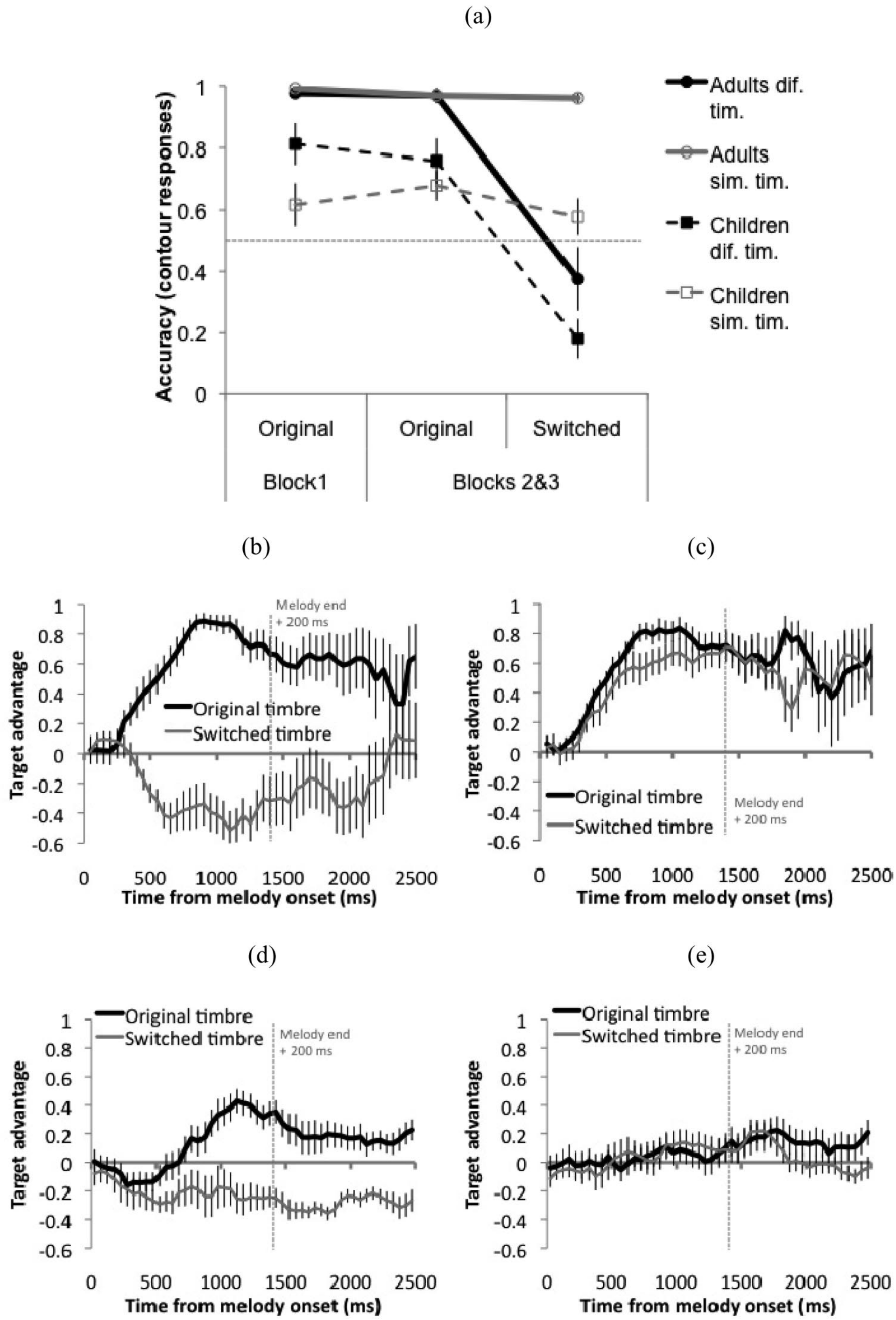


Figure 5. Experiment 2, (a) accuracy, and (b–e) looks to pictures in Blocks 2–3 for (b) adults, distinct timbres; (c) adults, similar timbres; (d) children, distinct timbres; (e) children, similar timbres. Error bars are standard errors.

significantly *below* chance accuracy on switched-timbre trials (.180 ± .258; $t(15) = 4.91$, $p = .0002$). Overall, child age marginally predicted accuracy, but only on original-timbre trials, $r(30) = .35$, $p = .05$.

Visual fixations. As before, trials with less than 50% of looks to pictures were dropped prior to analysis (13.9% of child trials, 18.0% of adults). ANOVAs with Timbre and Timbre Separation as factors assessed Target Advantage in Test Blocks 2–3 (Figure 5 b–e). Adults (Figure 5b–c) showed an effect of Timbre, $F(1, 28) = 39.65$, $p < .0001$, reflecting higher target advantage on original-timbre than switched-timbre trials. They also showed an effect of Timbre Separation, $F(1, 28) = 31.00$, $p < .0001$, with higher target advantage in the similar-timbres condition. Finally, Timbre and Timbre Separation interacted, $F(1, 28) = 39.65$, $p < .0001$. The interaction resulted from a null effect of Timbre for the similar-timbres condition ($F(1, 15) = 2.65$, $p = .12$; above chance overall, $t(15) = 9.54$, $p < .0001$), but a significant effect in the distinct-timbres condition, $F(1, 13) = 41.29$, $p < .0001$. In the distinct-timbres condition, target advantage was greater on original-timbre trials (above chance: $t(13) = 8.17$, $p < .0001$), than on switched-timbre trials, which showed *below*-chance target advantage: $t(13) = 3.46$, $p = .004$.

The results for children differed slightly. An effect of Timbre, $F(1, 30) = 5.55$, $p = .03$, reflected higher target advantage for original-timbre trials. There was no effect of Timbre Separation, but there was an interaction of Timbre and Timbre Separation, $F(1, 30) = 7.19$, $p = .01$. The interaction resulted from a null effect of Timbre in the similar-timbres condition ($F < 1$; target advantage at chance overall, $t(15) = 0.97$, $p = .35$), but a significant effect in the distinct-timbres condition, $F(1, 15) = 12.38$, $p = .003$. In the distinct-timbres condition, target advantage was greater on original-timbre trials (marginally above chance: $t(15) = 2.07$, $p = .06$), than on switched-timbre trials, which showed below-chance looks: $t(15) = 3.39$, $p = .004$. Interestingly, the only correlation with child age was with the switched-timbre trials, $r(30) = -.38$, $p = .03$. This negative correlation suggests that timbre-based responding was stronger as age increased.

Discussion

The results of the current experiment are somewhat parallel to Experiment 1. Adults showed sensitivity both to contour cues and to timbre cues: A distinct timbre difference swayed adults toward the timbre-matched response, but a mild timbre difference left their responses, and looks, unchanged. Children, however, appeared largely dependent on timbre. For a distinct timbre difference, their responses and looking patterns reversed completely. For a mild timbre difference, children's responses did not depend on timbre, but eye movements reflected less response clarity than in the distinct-timbres case. These results provide further suggestion that children are relatively insensitive to pitch contour.

But why are children so insensitive to pitch contour, given abundant evidence that much-younger listeners (infants) readily detect even more subtle contour changes (Trehub et al., 1984, 1985)? One possibility is that, by preschool, children have learned to discount pitch variation as an important cue in their environment, perhaps due to exposure to a non-tone language (English). Another possibility is that preschoolers, like infants, can discriminate contours easily, but have difficulty *encoding* contour information into long-term memory—and it would be impossible to succeed in the learning task in Experiments 1 and 2 without encoding melodies (and pictures, and melody-picture associations) into long-term memory.

The final experiment attempted to differentiate between these two alternatives—inattention to pitch variation versus good discrimination/poor long-term encoding. Experiment 3 tested children's abilities to identify changes to the contour, timbre, and pitch of melodies in an immediate same-different task, removing the requirement to encode melodies into long-term memory. If preschoolers simply ignore contour differences, then detection of contour changes should be at chance. However, if preschoolers are sensitive to contour changes but merely have difficulty with long-term encoding, then they should show good detection of contour changes.

Experiment 3

Method

Participants. N = 32 new 3- to 5-year-olds ($M = 4.4$ years, $SD = 0.5$, range: 3.5–5.2; 15 female) from the same population as before took part.

Stimuli. There were four exposure melodies and 24 additional test melodies. All melodies were either rising or falling, and falling melodies contained the additional timing cue as in previous experiments. The training melodies (Table 2, top) were arranged into four same and four different pairs, for use as examples and during reinforced training. Test melodies were arranged as specified in Table 2 to yield 20 same trials, and 24 different trials (eight different-contour, eight different-pitch, and eight different-timbre). The eight training pairs (four same, four different) were also mixed into the test to verify maintained adherence to task instructions. This yielded 48 test trials total.

Table 2
Stimuli in Experiment 3

Training trials	Stimuli
Same ($n = 4$) Different ($n = 4$)	Rising and falling sequences played by: tuba, C_3–G_3; harp, F#$_4$–C#$_5$. Different trials always differed in pitch height, timbre, and contour.
Test trials	**Stimuli**
Same, $n = 20$	Rising, falling; pitch ranges C_4–G_4, C_5–G_5 = 16 trials; all timbres plus tuba and harp from exposure = 4 trials.
Different contour, $n = 8$	Pitch range F#$_4$–C#$_5$; all timbres rising—falling (4 trials); falling—rising (4 trials).
Different pitch height, $n = 8$	All timbres octave separation (4 trials); half-octave separation (4 trials).
Different timbre, $n = 8$	Pitch range F#$_4$–C#$_5$; half rising, half falling similar timbres (bsn vs. sax); intermediate timbres (bsn vs. m tpt, sax vs. vib); different timbres (vib vs. m tpt).
All different, $n = 4$	These always differed in pitch height (C3–G3 vs. F#4–C#5), timbre (tuba vs. harp), and contour (rising vs. falling).

Note. Bsn = bassoon; sax = alto saxophone; m tpt = muted trumpet; vib = vibraphone.

Procedure. The task began with two visual example trials, one "different" (a circle and a triangle) and one "same" (two circles). Next were a *same* example trial using a high falling harp melody, which was repeated until children correctly answered "same" or accrued four repetitions; then, a *different* example trial (high falling harp melody vs. low rising tuba melody), which was repeated until children correctly answered "different" or accrued four repetitions. Next, children heard a block of eight clear same/different training trials with verbal feedback (Table 2, top). "Different" training trials always differed on all three dimensions: timbre (tuba or harp), pitch height (C_3–G_3 vs. F$\#_4$–C$\#_5$, a 1.5-octave disparity) and contour (rising vs. falling). Children had to answer correctly on at least seven of the eight training trials to continue. Four children failed to reach criterion (two after two training blocks, one after three blocks, one after five blocks). For all but two children (who performed similarly to the others), no more than three same or three different trials occurred in a row.

After passing the training criterion or failing to pass it after repeated training, children completed 48 test trials (see Table 2). Children who failed to reach the training criterion were included in analyses nonetheless, as there was no such basis for excluding children from the previous experiments. Thus, excluding them here might give an unduly advanced estimate of children's discrimination abilities.

Results

Overall accuracy was high (.738 ± .143). Children predominantly responded "same" on same trials (.825 ± .230), and "different" on the very-different trials from training (.930 ± .182), suggesting that they understood the task well. An ANOVA was conducted on e-logit transformed same responses (Figure 6a) to assess whether children reliably discriminated different contours, pitch heights, and timbres, with Trial Type (same, very-different, different contour, different pitch, different timbre) as a within-participants factor. Trial Type was significant, $F(4, 124) = 46.13$, $p < .0001$. Planned comparisons indicated that children responded "same" far less often for different-contour, different-pitch, and different-timbre trials (respectively, $t(31) = 6.47$, $p < .0001$; $t(31) = 7.74$, $p < .0001$; $t(31) = 9.14$, $p < .0001$) than for actual same trials (solid gray lines in Figure 6a). However, they did not discriminate perfectly: Children were reliably *more* likely to respond "same" to contour, pitch, and timbre differing pairs than to the highly different pairs ($t(31) = 3.96$, $p = .0004$; $t(31) = 5.97$, $p < .0001$; $t(31) = 3.97$, $p = .0004$; dashed gray lines in Figure 6a). Age did not significantly predict children's performance in any condition (all $p > .10$).

Within the critical trials, different timbres were dissociated more readily than different contours, $t(31) = 2.24$, $p = .03$ or different pitch heights, $t(31) = 3.32$, $p = .002$; solid black lines in Figure 6a. Planned comparisons examined the differing magnitudes of pitch height differences and timbre differences (Figure 6b). Octave pitch differences received fewer "same" responses than half-octave pitch differences, $t(31) = 3.82$, $p = .0006$. Distinct and intermediate timbres received fewer "same" responses than similar timbres, $t(31) = 2.82$, $p = .008$; $t(31) = 2.80$, $p = .009$.

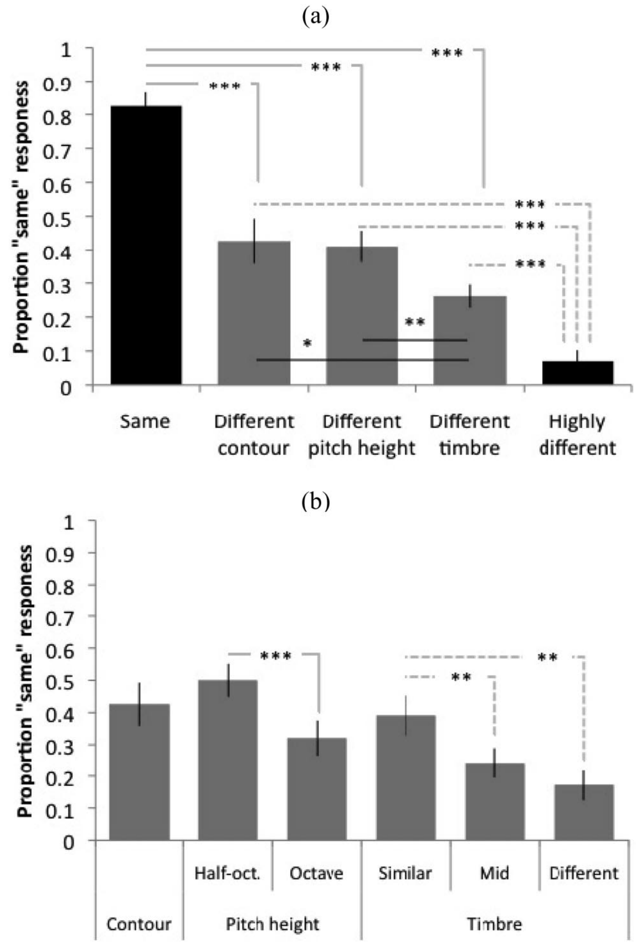


Figure 6. Experiment 3, (a) proportion "same" responses, (b) by trial type.

Discussion

Children reliably detected differences in contour, pitch height, and timbre between melodies like those used in Experiments 1–2. Their likelihood of detecting these changes related somewhat to ease in mapping these sound properties to characters in Experiments 1–2. That is, some of the changes that children detected more reliably in this discrimination task (timbre > contour) were the ones that they mapped to characters more readily. Nonetheless, even though children in Experiments 1–2 seemed to give little weight to contour in their memory representations, children in the current experiment detected contour differences fairly reliably. This suggests that contour, although discriminable, may be particularly difficult to encode into long-term memory.

General Discussion

The study asked whether pitch contour is central to preschool children's long-term representations of musical information. The answer appears to be no. In Experiment 1, children encoded pitch height more strongly than pitch contour. In Experiment 2, children encoded timbre more strongly than pitch contour. Adults were also influenced by pitch height and timbre cues, consistent with previous studies of adult sensitivity to pitch height (Creel & Tumlin, 2012; Schellenberg & Trehub, 2003) and timbre (Creel, 2012; Halpern & Müllensiefen, 2008; Radvansky, Fleming, & Simmons, 1995; Radvansky & Potter, 2000), but were additionally sensitive to pitch contour. Constrastingly, there was little evidence that children used pitch contour at all in the melody identification task. Nonetheless, children's *short-term* memory representations of pitch contour appear fairly accurate, as seen in Experiment 3's discrimination task. This suggests that pitch contour may be particularly difficult to encode in long(er)-term memory.

The three accounts posed initially were developmental stability, developmental shift, and developmental emergence. The results of the present study suggest a developmental shift for contour. To be more explicit, children barely use pitch contour, and adults use it robustly. Children appear more sensitive to pitch height and timbre, although there seems to be a quantitative change between children and adults' use of pitch height and timbre (adults were more accurate overall). It is not that children are more swayed by pitch height and timbre than adults were, but that they were *less* swayed by contour. These patterns are arguably most consistent with *developmental emergence*: some cues (pitch height, timbre) are available early and remain present, while others (contour) sharpen over experience. Additionally, the difficulty appears to be strongest in longer-term encoding; shorter-term memory representations examined in Experiment 3 appear more robust. The reasons why contour sensitivity might be slow to increase over development are discussed below.

One possible explanation for weak representations of contour is that contour changes are simply less discriminable than pitch height or timbre changes. If so, ease in the learning task should increase as the stimuli become more discriminable. On the surface this looks plausible. Children in the discrimination experiment (Experiment 3) showed less sensitivity to half-octave changes than to octave changes in pitch height, an effect seen in children's visual fixation data in Experiment 1; and, children in the discrimination experiment (Experiment 3) showed less sensitivity to contour changes than to timbre changes, and they were also less sensitive to contour than to timbre in Experiment 2. However, the pattern does not hold for pitch height versus contour: Children in Experiment 3 showed equivalent discrimination of changes in pitch height alone or contour alone, but pitch height cues outweighed contour cues in Experiment 1. This suggests that, tempting as it is to draw a connection, discriminability may not be the only factor which makes contour particularly difficult to represent in long-term memory.

Another possible explanation for weak use of contour in the learning task is that pitch contour degrades more rapidly in memory than do timbre and pitch height. Why might it degrade faster? This returns to the developmental emergence account: Perhaps contour is simply a richer domain which takes a greater amount of auditory-musical exposure to encode. That is, it has a much larger sample space than timbre or pitch height. Put in more concrete terms, pitch contour may have higher *information content* (it is less predictable; see, e.g., Pearce & Wiggins, 2006) than timbres or pitch heights. That is, each musical piece one hears has a fairly unique sequence of contours/relative pitches. Thus, although timbre and pitch height variability might be completely saturated by preschool age, the full space of contour possibilities is not.

Stability of Pitch Height and Timbre Information

The current study is consistent with previous research in suggesting stability of absolute pitch (or pitch height) representations (although note that adults here were more accurate than children in learning melodies separated by pitch height, suggesting some possible improvement; see Trehub et al., 2008). Like Stalinski and Schellenberg (2010), Experiment 1 suggests an early focus on pitch height, in an even younger age group (their youngest age tested was the oldest age tested here). The current study further expands on Stalinski and Schellenberg's work by demonstrating this pitch height focus in a paradigm with long-term memory demands. Also like Stalinski and Schellenberg, and like Creel and Tumlin (2012), the current work found absolute pitch sensitivity in adults as well. Interestingly, neither group showed any sensitivity to pitch *chroma* (the sense that 220 Hz and 440 Hz are both the note "A"). That is, the octave-pitch separation in Experiments 1 and 3 was treated as a starker change than the less-harmonically related half-octave pitch separation. This implies that memory representations of pitch height may not include chroma information, which is unlike classic absolute pitch in the pitch-labeling sense (e.g., Takeuchi & Hulse, 1993).

As a cautionary note, neither Stalinski and Schellenberg (2010) nor the current study distinguishes between true absolute pitch representations and what Creel and Tumlin (2012) called "global relative pitch." That is, listeners in both studies could be computing pitch height based not on absolute pitch but based on the pitches of the melody *relative to* one or more surrounding melodies. That is, there is always a pitch context present. Creel and Tumlin skirted this difficulty in a study on adults' melody recognition by changing the absolute pitches of their melodies after training, but maintaining the relative pitch relationships between melodies. They nonetheless found effects of the change in absolute pitch, overlaid on a strong effect of global relative pitch. If adults show some effects of absolute pitch, then presumably children would as well, but the degree to which they do so, and whether they do so more than adults, is unknown.

The current study is consistent with previous evidence of timbre sensitivity in young children (Trainor et al., 2004; Trehub et al., 1990). It is also consistent with timbre sensitivity in adults. For instance, adults recognize melodies (Halpern & Müllensiefen, 2008; Radvansky et al., 1995; Radvansky & Potter, 2000) and words (Goldinger, 1996, 1998; Palmeri, Goldinger, & Pisoni, 1993) better when those melodies are presented in their original timbre or when words are presented in their original voice. Adults also show some timbre-specificity in filling in metrical information (Creel, 2012). These studies considered together suggest that timbre sensitivity may be relatively stable, that is, that preschoolers may have timbre sensitivity similar to that of adults (see Vongpaisal, Trehub, & Schellenberg, 2009, for evidence that timbre aids young children in recognizing highly familiar TV show theme songs). Of course, the current study deliberately used two extremes of timbre similarity, and did not try to dissociate onset versus steady-state cues to timbre, so the exactness of the match in adult and child timbre sensitivity is still an open question.

Adding to the results of previous studies, the current results here are the first to show a high cue strength for timbre in children. That is, at least with a substantial timbre difference, timbre outweighs contour in melody recognition. Why is timbre such a strong cue?

It may be that timbre remains salient throughout life, as seen in several adult studies (Creel, 2012; Halpern & Müllensiefen, 2008; Radvansky et al., 1995; Radvansky & Potter, 2000). On the other hand, timbre may have an outsized influence in children because language guides them to focus on spectral characteristics. This language-tuning hypothesis is tantalizing, but its plausibility is hampered by the fact that language exposure steers children toward attending to spectral contrasts *in their native language* (e.g., Werker & Tees, 1984), not musical timbres. Nonetheless, it remains possible that attention to language (particularly a non-tone language like English) creates a more general bias toward spectral attending (see, e.g., Ladd et al., 2013).

Changes in Contour Processing

The strongest departure seen here from previous studies is children's apparent difficulty in encoding pitch contour. This is seemingly inconsistent with literature suggesting good processing of pitch contour in even younger children (as early as 5 months; Trehub et al., 1984), and sensitivity to contour and pitch changes in 4-year-olds (Corrigall & Trainor, 2010). However, all of those studies use either immediate discrimination paradigms (Stalinski & Schellenberg, 2010 with older children; Trehub et al., 1984, 1985 with infants), or highly familiar music (Corrigall & Trainor, 2010; Plantinga & Trainor, 2005; see below discussion on music familiarity). In the current study, preschool-aged children in an immediate discrimination task also performed well at detecting contour changes. Thus, the major difficulty seems to be *memory stability*.

One objection might be that, rather than reflecting weak long-term representations for contour, the learning task is simply "too difficult" for children. There are several counter-objections to this. First, it is not uniformly true that the task is difficult: the *learning task itself* is quite easy for children if they are given sufficiently distinguishable stimuli, a pattern that shows up in Experiment 2's distinct-timbre condition, and in other learning domains (voices; Creel & Jiménez, 2012). If the task were generally too difficult, then children would not be able to use any cue at all. Second, and more generally, one could counter that discrimination-type tasks are "too easy" in that they *over*estimate memory ability in infants and children. What counts as "too difficult" or "too easy" is essentially a question of what the more relevant measure is: immediate discrimination, or long-term recognition memory? If the question is whether infants and children have any sensitivity to perceptual differences whatsoever, then immediate discrimination paradigms are appropriate. If the question is how infants and children regularly process a wide range of incoming information, than recognition may be the more appropriate measure.

An additional objection might be that the music-referent mapping task used here underestimates children's music processing abilities, or provides an inaccurate picture of their weightings of cues, because it is not a naturalistic music task. A more naturalistic task might be simple recognition, that is, recognizing a song as familiar. This would remove the difficulty of association learning, and might perhaps be more like what children experience in the real world—recognizing a song as a thing in and of itself. If children were able to identify familiarized music based on contour after a long delay, this would suggest that association learning is the most difficult part of the task, rather than the memory durability account presented above. This would limit the results presented here to the situation of paired-association learning, rather than generalizing to all of music recognition. On the other hand, if children were poor at identifying familiarized music based on contour after a lengthy delay, this would suggest that memory durability, rather than association learning, presents the major difficulty, task naturalness notwithstanding. Future research should explore this possibility.

Familiarity and Contour Sensitivity

As noted above, greater contour and relative-pitch sensitivity has been obtained in research using familiar music (Corrigall & Trainor, 2010) or highly familiarized music (Plantinga & Trainor, 2005). This familiar–unfamiliar difference may be crucial, and plays back into the idea that contour/relative pitch predictability is relatively low (has a high information content; Pearce & Wiggins, 2006). Children may have isolated "islands" of predictability for familiar melodies that facilitate accurate processing of those melodies. This is similar to an argument advanced by Creel and Jiménez (2012) with regard to voice recognition. It is also highly similar to a set of findings in the developmental word recognition literature. Specifically, infants are more sensitive to small sound changes in familiar words (Swingley & Aslin, 2000, 2002) than in unfamiliar words (Stager & Werker, 1997)—even though much younger infants readily *discriminate* these sound changes. Preschoolers also show better recognition of familiar voices (Bartholomeus, 1973; Spence, Rollins, & Jerger, 2002) than of unfamiliar voices (Creel & Jiménez, 2012; Mann et al., 1979). These results together suggest that specific auditory familiarity boosts processing.

Why would there be a processing difference between familiar and unfamiliar words, or melodies? On a developmental emergence account, this may occur because young children's knowledge of speech sounds is inherent in, or closely bound to, their word representations. Similarly, in music learning, children's representations of musical pitch may be closely bound to their representations of particular melodies. Thus, when presented with unfamiliar melodies, their contour representations are noisy, because children are encoding unfamiliar melodies from scratch. By contrast, adults learn novel melodies (and novel words) through links to their sizable repertoire of musical (or verbal) listening knowledge.

Ongoing work should explore whether memory for contour fades more rapidly than memory for other musical attributes in children. This should be observable even in a discrimination task, if the intermelody interval is varied. With longer intervals, memory should begin to fade and result in less-accurate discrimination of a contour change.

Conclusion

Preschool-aged children show less sensitivity than adults to pitch contour in a music-referent mapping task. Nonetheless, children readily detect pitch contour changes in an immediate memory task. These results, taken together, suggest that (a) children have less memory stability for contour than for other musical attributes; (b) memory stability of contour increases over development, possibly due to its high information content; and (c) music development, like language development, may involve a protracted process of encoding detailed representations, referred to here as developmental emergence.

References

Barr, D. J. (2008). Analyzing “visual world” eyetracking data using multilevel logistic regression. *Journal of Memory and Language, 59,* 457–474. doi:10.1016/j.jml.2007.09.002

Bartholomeus, B. (1973). Voice identification by nursery school children. *Canadian Journal of Psychology, 27,* 464–472. doi:10.1037/h0082498

Bartlett, J. C., & Dowling, W. J. (1980). Recognition of transposed melodies: A key-distance effect in developmental perspective. *Journal of Experimental Psychology: Human Perception and Performance, 6,* 501–515. doi:10.1037/0096-1523.6.3.501

Boersma, P., & Weenink, D. (2010). *Praat: Doing phonetics by computer (Version 5.1.44)* [Computer program]. Retrieved 4 October 2010 from http://www.praat.org

Brainard, D. H. (1997). The psychophysics toolbox. *Spatial Vision, 10,* 433–436. doi:10.1163/156856897X00357

Cooper, R. M. (1974). The control of eye fixation by the meaning of spoken language. A new methodology for the real-time investigation of speech perception, memory, and language processing. *Cognitive Psychology, 6,* 84–107. doi:10.1016/0010-0285(74)90005-X

Cornelissen, F. W., Peters, E. M., & Palmer, J. (2002). The eyelink toolbox: Eye tracking with MATLAB and the psychophysics toolbox. *Behavior Research Methods, Instruments, & Computers, 34,* 613–617. doi:10.3758/BF03195489

Corrigall, K. A., & Trainor, L. J. (2010). Musical enculturation in preschool children: Acquisition of key and harmonic knowledge. *Music Perception, 28,* 195–200. doi:10.1525/mp.2010.28.2.195

Creel, S. C. (2012). Similarity-based restoration of metrical information: Different listening experiences result in different perceptual inferences. *Cognitive Psychology, 65,* 321–351. doi:10.1016/j.cogpsych.2012.04.004

Creel, S. C. (2014). Impossible to _gnore: Word-form inconsistency slows preschool children’s word-learning. *Language Learning and Development, 10,* 68–95. doi:10.1080/15475441.2013.803871

Creel, S. C., & Jiménez, S. R. (2012). Differences in talker recognition by preschoolers and adults. *Journal of Experimental Child Psychology, 113,* 487–509. doi:10.1016/j.jecp.2012.07.007

Creel, S. C., & Tumlin, M. A. (2012). Online recognition of music is influenced by relative and absolute pitch information. *Cognitive Science, 36,* 224–260. doi:10.1111/j.1551-6709.2011.01206.x

Goldinger, S. D. (1996). Words and voices: Episodic traces in spoken word identification and recognition memory. *Journal of Experimental Psychology: Learning, Memory, and Cognition, 22,* 1166–1183. doi:10.1037/0278-7393.22.5.1166

Goldinger, S. D. (1998). Echoes of echoes?: An episodic theory of lexical access. *Psychological Review, 105,* 251–279. doi:10.1037/0033-295X.105.2.251

Hallett, P. E. (1986). Eye movements. In K. R. Boff, L. Kaufman, & J. P. Thomas (Eds.), *Handbook of perception and human performance* (pp. 10.1–10.112). New York: Wiley.

Halpern, A. R., & Müllensiefen, D. (2008). Effects of timbre and tempo change on memory for music. *The Quarterly Journal of Experimental Psychology, 61,* 1371–1384. doi:10.1080/17470210701508038

Hannon, E. E., & Trehub, S. E. (2005a). Metrical categories in infancy and adulthood. *Psychological Science, 16,* 48–55. doi:10.1111/j.0956-7976.2005.00779.x

Hannon, E. E., & Trehub, S. E. (2005b). Tuning in to musical rhythms: Infants learn more readily than adults. *Proceedings of the National Academy of Sciences of the United States of America, 102,* 12639–12643. doi:10.1073/pnas.0504254102

Iverson, P., & Krumhansl, C. L. (1993). Isolating the dynamic attributes of musical timbre. *Journal of the Acoustical Society of America, 94,* 2595–2603. doi:10.1121/1.407371

Ladd, D. R., Turnbull, R., Browne, C., Caldwell-Harris, C., Ganushchak, L., Swoboda, K., . . . Dediu, D. (2013). Patterns of individual differences in the perception of missing-fundamental tones. *Journal of Experimental Psychology: Human Perception and Performance, 39*(5), 1386–1397. doi:10.1037/a0031261

Levitin, D. J. (1994). Absolute memory for musical pitch: Evidence from the production of learned melodies. *Perception & Psychophysics, 56,* 414–423. doi:10.3758/BF03206733

Mann, V. A., Diamond, R., & Carey, S. (1979). Development of voice recognition: Parallels with face recognition. *Journal of Experimental Child Psychology, 27,* 153–165. doi:10.1016/0022-0965(79)90067-5

Morton, J. B., & Trehub, S. E. (2001). Children’s understanding of emotion in speech. *Child Development, 72,* 834–843. doi:10.1111/1467-8624.00318

Navon, D. (1977). Forest before trees: The precedence of global features in visual perception. *Cognitive Psychology, 9,* 353–383. doi:10.1016/0010-0285(77)90012-3

Nelson, N. L., & Russell, J. A. (2011). Preschoolers’ use of dynamic facial, bodily, and vocal cues to emotion. *Journal of Experimental Child Psychology, 110,* 52–61. doi:10.1016/j.jecp.2011.03.014

Ohde, R. N., & Haley, K. L. (1997). Stop-consonant and vowel perception in 3- and 4-year-old children. *The Journal of the Acoustical Society of America, 102,* 3711–3722. doi:10.1121/1.420135

Palmeri, T. J., Goldinger, S. D., & Pisoni, D. B. (1993). Episodic encoding of voice attributes and recognition memory for spoken words. *Journal of Experimental Psychology: Learning, Memory, and Cognition, 19,* 309–328. doi:10.1037/0278-7393.19.2.309

Pearce, M. T., & Wiggins, G. A. (2006). Expectation in melody: The influence of context and learning. *Music Perception, 23,* 377–405. doi:10.1525/mp.2006.23.5.377

Pelli, D. G. (1997). The VideoToolbox software for visual psychophysics: Transforming numbers into movies. *Spatial Vision, 10,* 437–442. doi:10.1163/156856897X00366

Plantinga, J., & Trainor, L. J. (2005). Memory for melody: Infants use a relative pitch code. *Cognition, 98*(1), 1–11. doi:10.1016/j.cognition.2004.09.008

Quam, C., & Swingley, D. (2010). Phonological knowledge guides two-year-olds’ and adults’ interpretation of salient pitch contours in word learning. *Journal of Memory and Language, 62,* 135–150. doi:10.1016/j.jml.2009.09.003

Quam, C., & Swingley, D. (2012). Development in children’s interpretation of pitch cues to emotions. *Child Development, 83,* 236–250. doi:10.1111/j.1467-8624.2011.01700.x

Radvansky, G. A., Fleming, K. J., & Simmons, J. A. (1995). Timbre reliance in nonmusicians’ and musicians’ memory for melodies. *Music Perception, 13,* 127–140. doi:10.2307/40285691

Radvansky, G. A., & Potter, J. K. (2000). Source cuing: Memory for melodies. *Memory & Cognition, 28,* 693–699. doi:10.3758/BF03198403

Saffran, J. R., & Griepentrog, G. J. (2001). Absolute pitch in infant auditory learning: Evidence for developmental reorganization. *Developmental Psychology, 37,* 74–85. doi:10.1037/0012-1649.37.1.74

Schellenberg, E. G., & Trehub, S. E. (1996). Natural musical intervals: Evidence from infant listeners. *Psychological Science, 7,* 272–277. doi:10.1111/j.1467-9280.1996.tb00373.x

Schellenberg, E. G., & Trehub, S. E. (2003). Good pitch memory is widespread. *Psychological Science, 14,* 262–266. doi:10.1111/1467-9280.03432

Sergeant, D. (1969). Experimental investigation of absolute pitch. *Journal of Research in Music Education, 17,* 135–143. doi:10.2307/3344200

Sergeant, D., & Roche, S. (1973). Perceptual shifts in the auditory information processing of young children. *Psychology of Music, 1,* 39–48. doi:10.1177/030573567312006

Smith, N. A., & Schmuckler, M. A. (2008). Dial A440 for absolute pitch: Absolute pitch memory by non-absolute pitch possessors. *Journal of the Acoustical Society of America, 123,* EL77–EL84. doi:10.1121/1.2896106

Spence, M. J., Rollins, P. R., & Jerger, S. (2002). Children's recognition of cartoon voices. *Journal of Speech, Language, and Hearing Research, 45,* 214–222. doi:10.1044/1092-4388(2002/016)

Stager, C. L., & Werker, J. F. (1997). Infants listen for more phonetic detail in speech perception than in word-learning tasks. *Nature, 388,* 381–382. doi:10.1038/41102

Stalinski, S. M., & Schellenberg, E. G. (2010). Shifting perceptions: Developmental changes in judgments of melodic similarity. *Developmental Psychology, 46,* 1799–1803. doi:10.1037/a0020658

Swingley, D., & Aslin, R. N. (2000). Spoken word recognition and lexical representation in very young children. *Cognition, 76,* 147–166. doi:10.1016/S0010-0277(00)00081-0

Swingley, D., & Aslin, R. N. (2002). Lexical neighborhoods and the word-form representations of 14-month-olds. *Psychological Science, 13,* 480–484. doi:10.1111/1467-9280.00485

Takeuchi, A. H., & Hulse, S. H. (1993). Absolute pitch. *Psychological Bulletin, 113,* 345–361. doi:10.1037/0033-2909.113.2.345

Tanenhaus, M. K., Spivey-Knowlton, M. J., Eberhard, K. M., & Sedivy, J. C. (1995). Integration of visual and linguistic information in spoken language comprehension. *Science, 268,* 1632–1634. doi:10.1126/science.7777863

Trainor, L. J. (2005). Are there critical periods for musical development? *Developmental Psychobiology, 46,* 262–278. doi:10.1002/dev.20059

Trainor, L. J., Wu, L., & Tsang, C. D. (2004). Long-term memory for music: Infants remember tempo and timbre. *Developmental Science, 7,* 289–296. doi:10.1111/j.1467-7687.2004.00348.x

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

Trehub, S. E., Endman, W., & Thorpe, L. A. (1990). Infants' perception of timbre: Classification of complex tones by spectral structure. *Journal of Experimental Child Psychology, 49,* 300–313. doi:10.1016/0022-0965(90)90060-L

Trehub, S. E., Schellenberg, E. G., & Nakata, T. (2008). Cross-cultural perspectives on pitch memory. *Journal of Experimental Child Psychology, 100,* 40–52. doi:10.1016/j.jecp.2008.01.007

Trehub, S. E., Thorpe, L. A., & Morrongiello, B. A. (1985). Infants' perception of melodies: Changes in a single tone. *Infant Behavior and Development, 8,* 213–223. doi:10.1016/S0163-6383(85)80007-2

Trueswell, J. C., Sekerina, I., Hill, N. M., & Logrip, M. L. (1999). The kindergarten-path effect: Studying on-line sentence processing in young children. *Cognition, 73,* 89–134. doi:10.1016/S0010-0277(99)00032-3

Vongpaisal, T., Trehub, S. E., & Schellenberg, E. G. (2009). Identification of TV tunes by children with cochlear implants. *Music Perception, 27,* 17–24. doi:10.1525/mp.2009.27.1.17

Ward, W. D. (1999). Absolute pitch. In D. Deutsch (Ed.), *Psychology of music* (2nd ed., pp. 265–298). New York, NY: Academic Press. doi:10.1016/B978-012213564-4/50009-3

Weiss, M. W., Trehub, S. E., & Schellenberg, E. G. (2012). Something in the way she sings: Enhanced memory for vocal melodies. *Psychological Science, 23,* 1074–1078. doi:10.1177/0956797612442552

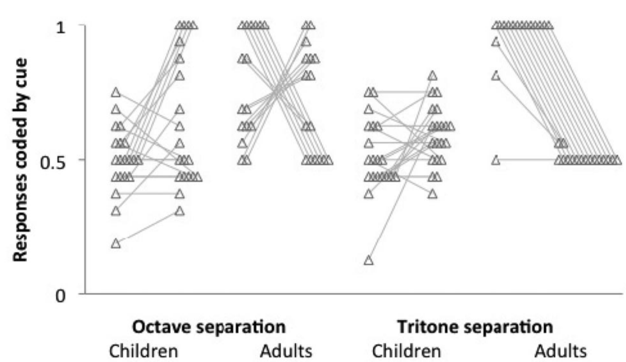
Werker, J. F., & Tees, R. (1984). Cross-language speech perception: Evidence for perceptual reorganization during the first year of life. *Infant Behavior and Development, 7,* 49–63. doi:10.1016/S0163-6383(84)80022-3

Appendix

Individual Response Patterns in Experiments 1 and 2

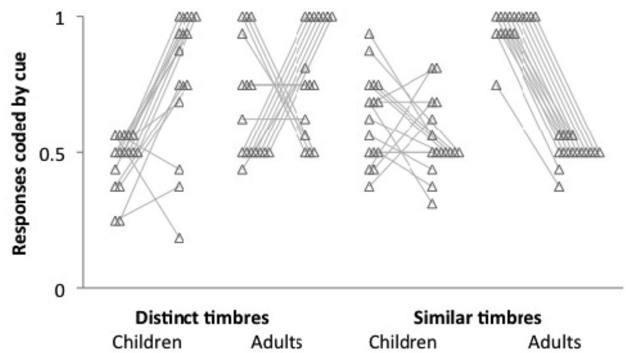
The left end of each line denotes each individual's accuracy in terms of contour; the right end denotes each individual's accuracy in terms of pitch height (top figure) or timbre (bottom figure). 1 = responded perfectly based on that cue; 0.5 = responded at chance (overall) based on that cue.

Experiment 1



(*Appendix continues*)

Experiment 2



Received August 20, 2013
Revision received January 13, 2014
Accepted January 15, 2014 ■