Title

The familiar-melody advantage in auditory perceptual development: Parallels between spoken language acquisition and general auditory perception

Journal
Attention, Perception, & Psychophysics, 81(4)

ISSN
1943-3921

Author
Creel, Sarah C

Publication Date
2019-05-01

DOI
10.3758/s13414-018-01663-7

Peer reviewed

The familiar-melody advantage in auditory perceptual development:

Parallels between spoken language acquisition and general auditory perception

Sarah C. Creel

Cognitive Science

UC San Diego

9500 Gilman Drive Mail Code 0515

La Jolla, CA 92093

Ph: 858.534.7308

Email: screel@ucsd.edu

Abstract

How do learners build up auditory pattern knowledge? Findings from children's spoken word learning suggest more robust auditory representations for highly-familiar words than for newly-learned words. This argues against spoken language learning as a process of simply acquiring a fixed set of speech sound categories, suggesting that specific words may be the relevant units. More generally, one might state this as the *specific learning hypothesis*—that acquiring sound pattern knowledge involves learning specific patterns, rather than abstract pattern components. To understand the nature of human language knowledge, it is important to determine whether this specific learning reflects processes unique to spoken language learning, or instead reflects more general auditory learning processes. To ask whether the specific learning hypothesis extends to auditory pattern learning more generally, the current study tested perceptual processing of familiar melodies vs. carefully-matched unfamiliar melodies. Children performed better at both audiovisual mapping (Experiment 1) and same-different auditory discrimination (Experiment 2) when hearing familiar melodies than when hearing unfamiliar melodies. This is consistent with the specific learning hypothesis and with exemplar-style general-auditory accounts of pattern learning, though alternative explanations are possible. *183 words*

A major question in speech processing, and in spoken language development, is the extent to which auditory representations of language depend on neurally prespecified sound patterns restricted to the domain of language, or instead represent the outcome of more general learning principles. Diehl and colleagues (Diehl & Walsh, 1989; Hay & Diehl, 2007; Kluender, Diehl, & Wright, 1988) have written extensively about overlap between speech processing and general auditory processes. Their research suggests that speech sound patterns depend on or emerge from general auditory learning processes, and that speech sound patterns do not depend on having prespecified categories (see especially Kluender, Diehl, & Killeen, 1987, who showed speech sound learning in quail). On this second account, one should see parallels between auditory pattern learning in speech and in nonspeech.

In this spirit, the current work explores learning of sound patterns of nonspeech auditory stimuli. To further substantiate an account of spoken language learning in terms of general auditory processes, we ask whether nonspeech auditory learning shows analogous pattern learning processes. We explore an *exemplar view* on the learning process. On exemplar accounts (e.g. Goldinger, 1996), representations emerge from accruing many specific exemplars, or large numbers of neural traces. Only after collecting a large number of representations do broader-scale patterns—such as speech sounds apart from particular word contexts, or generic properties of musical scales—emerge. This suggests that early in the learning process, for example, in childhood, learners may distinguish *well-known* sound patterns that differ subtly, while failing to distinguish *novel patterns* that differ by the same characteristics.

It is worth contrasting exemplar accounts with other explanatory frameworks, such as prototype theories of memory formation. On an exemplar account, recognition is computed by a composite of traces which are each activated in proportion to their similarity to the input. On a prototype account, recognition is accomplished by comparing a new instance to each category's central tendency. Thus, both types of accounts are probabilistic. One area where they differ, though, is that many prototype or prototype-like accounts of pattern learning specify a low-level unit of analysis in terms of speech sounds (e.g. Kuhl, Williams, Lacerda, Stevens, & Lindblom, 1992), or, in music, in terms of properties like musical key membership (e.g. Trainor & Trehub, 1994). By contrast, exemplar accounts often take events, such as entire words, as the relevant unit of analysis (see, e.g., Goldinger, 1996). A prototype account that posits melodies as the unit of analysis might make similar predictions in the present study. Still, exemplar accounts also have other appealing properties in that the perceptual category does not need to be pre-assigned, and can instead emerge, unsupervised, from the input; and that category variability is implicitly preserved. However, in the current case, the appeal of exemplar approaches is less about their exact similarity computations as much as what the unit of analysis is.

Empirical support for exemplar-like auditory event memory

One set of findings that suggest auditory event memories consist of specific exemplars comes from the domain of early child word learning. Infants (14-month-olds) can distinguish similar speech sounds such as /b/ and /d/ in immediate discrimination (detecting a change to "dih" after many repetitions of "bih") but they perform poorly at learning labels distinguished by these words ("bih" vs. "dih") at 14 months, only succeeding at 17-20 months (Stager & Werker, 1997; Werker, Fennell, Corcoran, & Stager, 2002). One interpretation of this disconnect between discrimination and word learning is that word learning requires maintaining longer-term sound-pattern representations than discrimination, which requires only short-term representations. However, contrasting with novel-word learning findings, *familiar* words that are distinguished by these speech sounds (e.g. *ball* vs. *doll*) are readily recognized by children around 14 months (Fennell & Werker, 2003; Swingley & Aslin, 2002). Note that the critical contrast here is between two similar studies, rather than within a single study. On the one hand, Stager and Werker (1997) presented children with two unfamiliar pictures, each with an unfamiliar label—*bih* or *dih*—and then tested their learning of the labels. Fennell and Werker (2003) used exactly the same procedure, and exactly the same b/d speech sound contrast, but used instead the familiar words *ball* and *doll* and pictures of a ball and a doll. While the *bih*/*dih* children failed at 14 months (Stager & Werker, 1997), the *ball*/*doll* children succeeded (Fennell & Werker, 2003), indicating that they could use the b/d contrast to tell words apart, but only when embedded in familiar words. This suggests that children may not yet have separable representations of individual speech sounds, but instead have representations of entire words, consistent with an exemplar account.

Music perception and specific auditory memory

There is evidence of a similar phenomenon in the music perception literature. For example, McFadden and Callaway (1999) found, in a series of psychophysical experiments, that adult listeners were more sensitive to subtle changes in familiar musical (and speech) materials than in matched unfamiliar materials. Evidence from child music perception is suggestive but not conclusive. Research by Trainor and colleagues suggests that 4-year-olds are less perceptually sensitive than adults to musical structure violations for unfamiliar melodies (Trainor & Trehub, 1994), but are excellent at detecting musical structure violations in familiar melodies, specifically, Twinkle Twinkle Little Star (Corrigall & Trainor, 2010). In a related finding which mirrors the Stager and Werker (1997) result of discrimination ability without word-learning ability, Creel (2014b, 2016) reported that preschool-aged children can auditorily discriminate some musical patterns (rising vs. falling pitches) that they cannot associate with visual objects (cartoons; Creel, 2014b, 2016; see also Pajak, Creel, & Levy, 2016, for a similar effect in adult second-language word learning). This is especially striking in that the hard-to-associate sounds differ in *pitch contour*, a musical property to which even young infants are sensitive (Trehub, Bull, & Thorpe, 1984),[1] and one regarded as central to melodic identity (Dowling, 1978). Of course, Creel's (2014b, 2016) audiovisual association tasks used *unfamiliar* brief melodies, raising the possibility that more familiar melodies are more associable.

These musical and linguistic findings seem at odds with accounts of learning as a process of acquiring speech sound categories (Werker & Tees, 1984) or acquiring abstract musical knowledge such as scale structure (Lynch, Eilers, Oller, & Urbano, 1990). They are easier to square with an account that children are accruing specific exemplars of auditory patterns, such as melodies and words, rather than directly learning more abstract structures (major scale, phonology). Thus, relatively weaker performance with less-familiar materials is driven by weaker underlying auditory memory representations (of the unfamiliar nonsense word *bih* or an unfamiliar melody) relative to more-familiar representations (of the familiar word *ball* or a familiar melody like Twinkle Twinkle Little Star). On this **specific learning hypothesis**, more-

---

[1] *Creel (2014b, 2016; see also Creel & Quam, 2015) have attributed this seeming developmental disconnect—good performance by infants but some difficulties in preschoolers—to substantial differences in infant vs. older-child test paradigms. One cannot perform a habituation test with a 4-year-old, nor can one conduct a same-different test with a 14-month-old.*

familiar items should be both *easier to discriminate* and *easier to associate* than unfamiliar items with similar properties. However, the most supportive evidence for this hypothesis comes from the word learning literature, and even that evidence is scattered across multiple studies.

The current study

To assess whether the specific-learning hypothesis applies to nonspeech auditory patterns as well as spoken language, the current study tested preschool-aged children's abilities to form audiovisual associations with, and to discriminate, either familiar melodies or scrambled (thus unfamiliar) versions of those melodies. If children perform with equal accuracy on familiar vs. unfamiliar songs, this would fail to support the specific learning hypothesis. It would also suggest that musical pattern learning operates differently from pattern learning in language, inconsistent with a general-auditory account of pattern learning. On the other hand, better performance on familiar songs than unfamiliar songs would lend credence to the specific-learning hypothesis. It would also support a general-auditory view of pattern acquisition. That is, if we find familiar-form effects in nonspeech auditory processing similar to those found previously in spoken word processing, then it suggests that pattern-learning processes are general across multiple types of auditory events, rather than being isolated to spoken language.

Pretesting

Initial pretesting aimed to determine what melodies children were best at identifying, out of a set of likely candidates. The first pretest tested 19 preschool-aged children by presenting six childhood melodies (Mary Had a Little Lamb, Twinkle Twinkle Little Star, Happy Birthday, Deck the Halls, Yankee Doodle, and London Bridge), and asking children to name them. After this, melodies will be referred to with only the first word of the title for brevity. In the first block of trials, two measures (6-8 notes) of each song were presented once; in the second block, four measures (12-17 notes) of each song were presented once. Melodies were played in a synthesized piano timbre. Accuracy was defined as naming any semantic content in the song, not just the title or lyrics; for example, for “Mary Had a Little Lamb,” the answers “Mary”, “lamb,” and “sheep” were all accepted. Even by this lenient criterion, naming accuracy (Table 1) was fairly low (21% overall) and did not differ between 2-measure and 4-measure melodies. The highest naming rates were for Happy (36%), Mary (33%), and Twinkle (39%). However, since children may possess some perceptual knowledge of the songs but have difficulty verbalizing it, we conducted a second pretest.

The second pretest with 23 children used a 2-alternative forced-choice picture-matching task with the same melodies, except than London Bridge (0% naming recognition) was replaced with the Star-Spangled Banner. Pictures were as described in Table 1. Here, overall accuracy was 59%. Length (2 vs. 4 measures) did not impact recognition. The pairs with highest accuracy were Mary vs. Twinkle and Birthday vs. Twinkle (both 67%; note that Table 1 shows per-melody accuracy). Since Mary and Twinkle have identical rhythmic patterns, we selected this melody pair, allowing us to examining pitch contour effects in isolation from timing differences. Children’s modest pretest performance on familiar songs is revisited in the General Discussion.

As described above, previous studies (Creel 2014b, 2016) have shown that children have difficulty distinguishing musical patterns by their pitch contours. Thus, one option for the current study was to test these two familiar melodies and compare our results qualitatively to previous studies. This parallels work in child language research (Fennell & Werker, 2003; Stager & Werker, 1997) and child music research (Trainor & Trehub, 1994; Corrigall & Trainor, 2010), where findings of sensitivity to familiar vs. unfamiliar sound patterns are not compared in the same study. However, comparing the two familiar songs here with unfamiliar songs from previous studies was less than desirable, because the melodies used in earlier work were shorter and less complex (4-5 notes, unidirectional pitch contours) than melodies used here (7 notes, and bidirectional pitch contours). A better unfamiliar control, then, would be melodies with properties more directly matched to the familiar melodies used. Therefore, after selecting the two familiar melodies, we created scrambled-note versions of each (Figure 1). (Note that the note-scrambling was done at the notation level, so that we were not actually rearranging segments of an audio file, which could lead to unnatural juxtapositions of reverberation from previous notes.) We used simple contours for each, so that familiar and unfamiliar melodies were matched for overall naturalness. Both scrambled melodies ended on the same note as their original melodies, so that note duration was realistic (final notes of phrases tend to be longer) and so that the degree of key resolution was similar to the originals. This use of control unfamiliar stimuli within the same study represents an advance over previous studies in both the word recognition and music perception literatures.

Table 1. Pretest stimuli.

Melody	**Picture**	**Naming accuracy**	**Recognition accuracy**
Mary Had a Little Lamb	Lamb	0.325	0.612
Twinkle Twinkle Little Star	Star	0.390	0.670
Happy Birthday	cake with candles	0.361	0.627
Deck the Halls	Christmas tree	0.182	0.533
Yankee Doodle	US Revolutionary soldier	0.000	0.536
London Bridge	(not used in pretest 2)	0.000	N/A

Star-Spangled Banner	Flag	N/A	0.565

Experiment 1: association learning

The first experiment asked children to associate two melodies with two different pictures. Since previous studies (Creel, 2016) suggested that children were unable to associate two melodies differing only in contour with two different pictured objects, we reasoned that the strongest test of melody familiarity effects would come from this difficult task. We were also interested in whether children would use preexisting associations with familiar melodies, or if familiar-melody benefits extended to both melody-related and novel pictures. Therefore, half of children were asked to learn that melody-related pictures "go with" familiar or unfamiliar melodies, while the other half were asked to learn that *novel* pictures—the cartoon characters used in previous melody-picture association studies (Creel, 2014b, 2016)—"go with" familiar or unfamiliar melodies.

Method

Participants. Sixty-four preschool-aged children (ages 3-5) took part. A second replication sample of 64 children was also obtained. Results are reported together for clarity and brevity. Five additional children were excluded from analysis because they did not complete the task.

Stimuli. The four melodies were Mary, Twinkle, Mary-scrambled, and Twinkle-scrambled (Figure 1). Melodies were synthesized in piano timbre in Finale 2009 software (MakeMusic, Inc.) and were edited and scaled to 70 dB SPL in Praat (Boersma & Weenink, 2014). Two different picture sets were used (Figure 2): either a lamb and a star (familiar-picture condition), or two cartoon characters (unfamiliar-picture condition). Related pictures were chosen to be related to the lyrics of each song, on analogy with Fennell and Werker's (2003) study using the words *ball* and *doll* and pictures of those objects. The two cartoon characters have been used in a variety of studies of word learning (Creel, 2014a, 2014c), talker-voice learning (Creel & Jiménez, 2012), and melody learning (Creel, 2014b, 2016). In all cases where learned elements were perceptually distinct, children achieved high rates of learning accuracy (80%+). Thus, if children have difficulty associating melodies with the cartoon characters, it is not due to difficulty visually discriminating the cartoon characters.

Procedure. Children received reinforced learning trials in blocks of 8 trials each (4 with one melody-picture combination, 4 with the other). In both the original and replication studies, 16 children each learned to associate familiar melodies with related pictures; scrambled melodies with related pictures; familiar melodies with novel pictures; scrambled melodies with novel pictures. In related-picture conditions, children were told that they would see a star sing the star song, and a lamb singing the lamb song. In novel-picture conditions, children were told that they would see one creature sing the star song, and another creature sing the lamb song. On each trial, two pictures appeared on the left and the right side of the screen (side counterbalanced across trials), and then a melody played. The child was asked to select the creature who sang the song. Once they scored at least 7/8 in a block of learning trials, or completed five learning blocks, they continued to unreinforced test trials.

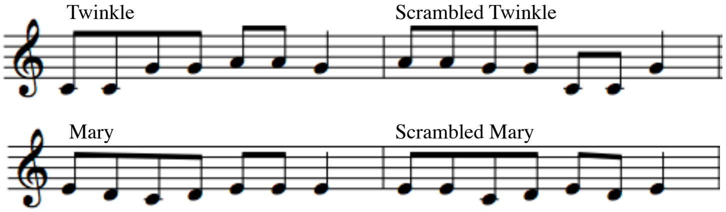

Figure 1. Experiments 1 and 2, melody stimuli.

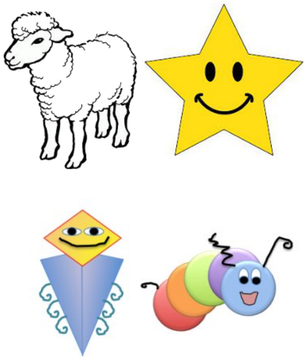

Figure 2. Experiment 1, visual stimuli. Top row: related pictures. Bottom row: novel pictures.

Results

To analyze data (Figure 3), we used a logistic regression model in R (R Core Team, 2016) using the lme4 package (Bates, Maechler, Bolker, & Walker, 2015). Logistic regression takes into account the fact that accuracy is binomially distributed rather than normally distributed. The dependent variable was accuracy, with correct coded as 1 and incorrect coded as 0. This model used the predictors Melody Familiarity (familiar, scrambled), Picture Familiarity (related pictures, novel cartoons), Feedback (reinforced learning trials, nonreinforced test trials), and their interactions. To account for the within-subjects nature of the data, the model included random intercepts for participants, and Feedback random slopes for participants (other variables were between-subjects). For feedback trials, results from just the first block of learning trials—the only block where all participants took part—were analyzed.

Original sample. There was an effect of Feedback (B=0.16, SE=0.06, *z*=2.70, *p*=.007), suggesting that children became less accurate once reinforcement was no longer available. There was also a Melody Familiarity x Picture Familiarity interaction (B=0.30, SE=0.08, *z*=3.53, *p*=.0004), apparently stemming from an advantage for familiar over scrambled melodies which occurred only when pictures were related. However, this interaction was qualified by a 3-way Feedback x Melody Familiarity x Picture Familiarity interaction (B=0.13, SE=0.06, *z*=2.35, *p*=.02). To examine this 3-way interaction, we examined the simple 2-way interactions of Melody Familiarity x Picture Familiarity, separately for reinforced and nonreinforced trials. For reinforced (learning) trials, the effect of Melody Familiarity was significant (B=0.24, SE=0.12, *z*=2.03, *p*=.04), as was the 2-way interaction (B=0.47, SE=0.12, *z*=3.92, *p*<.0001). Examining

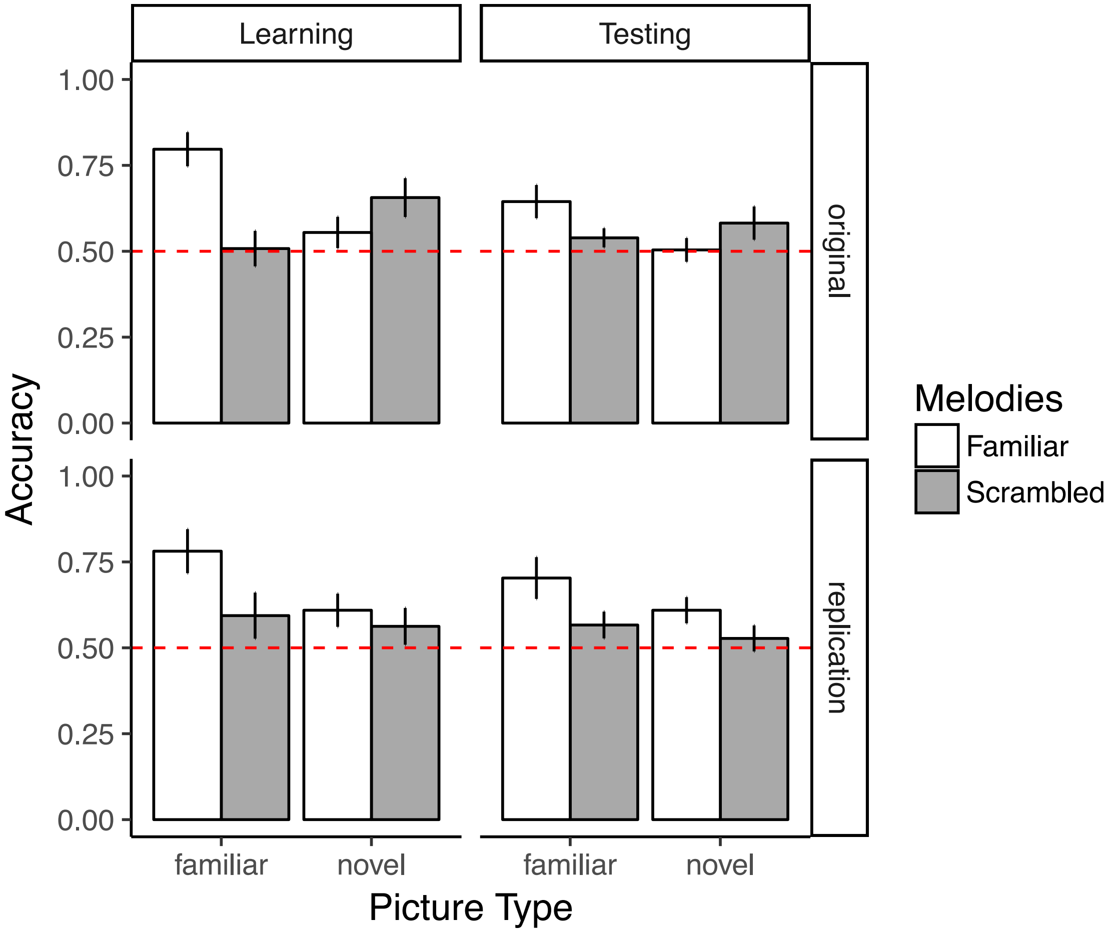

Figure 3. Experiment 1 (upper) and its replication (lower), accuracy in first block of learning trials (left) and in test trials (right). Dashed line represents chance performance. Error bars are standard errors. Plots created in R using ggplot2 (Wickham, 2016).

Melody Familiarity at each level of Picture Familiarity revealed that only when the picture was related was there a significant benefit (B=0.73, SE=0.19, z=3.86, p=.0001) of familiar melodies over scrambled ones.

We then looked at the Melody Familiarity x Picture Familiarity interaction for nonreinforced trials. There was no main effect of Melody Familiarity, but the interaction was significant (B=0.20, SE=0.09, z=2.34, p=.02). Looking at each level of Picture Familiarity separately, the effect of Melody Familiarity was marginally significant (B=0.23, SE=0.12, z=1.94, p=.051) for related pictures, but not significant for novel pictures. No other effects were significant. To restate the interaction pattern: when pictures are related, children are better at associating familiar melodies than at associating scrambled melodies. For novel melodies, there is no such familiar-melody advantage. The three-way interaction with Feedback appears to come from the fact that the 2-way interaction is larger in magnitude for the learning (reinforced) trials.

Replication sample. The same analysis was performed on the replication sample. The effect of Feedback was marginally significant (B=0.12, SE=0.07, z=1.68, p=.09), with slightly lower performance when reinforcement ceased. There was an effect of Picture Familiarity (B=0.22, SE=0.10, z=2.13, p=.03), with higher accuracy for related pictures than novel pictures. There was also an effect of Melody Familiarity (B=0.30, SE=0.10, z=2.91, p=.004), with higher accuracy for familiar melodies than novel melodies. No interactions were significant. However, because of the asymmetric effects of Melody Familiarity in the original sample, we examined Melody Familiarity effects at each level of Picture Familiarity. For related pictures, the effect of Melody Familiarity was significant (B=0.47, SE=0.18, z=2.56, p=.01), with stronger performance on familiar melodies. However, for novel pictures, the effect of Melody Familiarity was not significant.

Discussion

Both the original experiment and the replication suggest that melody familiarity has strong effects on melody encoding, mainly when the association naturally fits with preexisting song knowledge. This closely mirrors effects found in the word-learning literature, where 14-month-old children respond to slight differences in pronunciation of familiar words (*ball* vs. *doll*; Fennell & Werker, 2003) but do not respond to slight differences in pronunciations of newly-learned words (*bih* vs. *dih*; Stager & Werker, 1997). We also tested an additional question, of whether the familiar-sound-pattern advantage extends to novel associations. This has not been addressed in the word-learning literature (such as labeling two novel objects "ball" and "doll"). What we find here is that the familiar-melody advantage may not extend to novel picture associations. This is especially interesting in that the instructions pointed out the connection to children even for novel pictures: they were explicitly told that one creature sang the "star song" and the other sang the "lamb song." It seems in principle that they could simply have remembered one creature as the star one, and the other as the lamb one, yet they did not seem to do so.

Of course, in both the current study and possible future child word-learning studies, pre-existing associations might interfere with learning new ones. That is, knowing what "ball" refers to, or having preexisting semantic associations with Twinkle Twinkle Little Star, might interfere with associating each with new, unfamiliar visual materials. This appears to be the case more so in the original experiment, where novel pictures show a numerically reversed melody-familiarity effect, than in the replication, where they do not. Given the variability between these two patterns of results, we decline to make a strong statement. It is worth noting that in the child word-learning literature, homophone interference effects (inability to learn that an unfamiliar object is called a "rope") appear to surface mainly when a rope is present as a response choice (Doherty, 2004). Thus, concerns of existing-association interference may be strongest in cases where the existing associate is present to interfere. Further, Storkel and Maekawa (2005) report *better* naming accuracy when children have learned that a novel object has a homophonous name like "comb," vs. nonsense-word name like "bine," which is consistent with a sound-pattern familiarity effect in word learning.

In any case, finding here suggest that, unlike previous studies which used only unfamiliar melodies (Creel, 2014b, 2016), children can learn melodic contour-picture associations, and do so better for familiar than unfamiliar melodies. It also raises questions as to whether the familiarity advantage is at the level of perceptual representation, at the level of learned associations (between musical patterns and lyrics or contexts), or both. Therefore, the next experiment tested whether children perform better on familiar melodies in a task that more directly assesses perceptual processing: a same-different task. If children have more robust perceptual representations for familiar than for unfamiliar (scrambled) melodies, then discrimination performance should be better for changes from one familiar melody to another than for changes from one scrambled melody to another.

The study allowed an additional test of potentially heightened salience for familiar melodies. Previous studies of brief tone sequence discrimination have suggested that children are less sensitive to melodic contour than to timbre, that is, musical-instrument sound quality (Creel, 2014b, 2016). Therefore, the next experiment also included timbre discrimination trials, at three different levels of timbre similarity. This allowed us to assess whether differences between *familiar* melodies (vs. the unfamiliar ones, as tested previously) might be *more* salient than timbre differences.

Experiment 2: discrimination

Method

Participants. Ninety preschool-aged children (ages 3-5 years) who had not taken part in Experiment 1 participated. An additional 51 children took part but were excluded due to: failure to meet training criterion (29; see Creel, Weng, Fu, Heyman, & Lee, 2018, for similar rates of failure to meet training criterion), noise disruptions at testing site (5), failure to complete the study (5), unwillingness or inability to follow instructions (2), extreme inattentiveness (1), computer errors (3), or having exposure to a tone language (6), which may change performance on the task (Creel et al., 2018).

Stimuli. Melodies were synthesized in multiple timbres in Finale 2009 software (MakeMusic, Inc.) and were edited and scaled to 70 dB SPL in Praat (Boersma & Weenink, 2014). Mary, Twinkle, Mary-scrambled and Twinkle-scrambled were heard by all participants. Participants also heard different-timbre trials where the same melody was played but the timbre changed. For 1/3 of participants each, the two timbres were very similar (piano, guitar; see Fragoulis, Papaodysseus, Exarhos, Roussopoulos, Panagopoulos, & Kamarotos, 2006) moderately similar (bassoon and alto saxophone); or distinct (vibraphone and muted trumpet). Timbre distinctiveness estimates for the latter two pairs were drawn from Iverson and Krumhansl (1993). For a child in a given timbre condition, the different-contour trials occurred equally often in the two timbres used in that experiment. That is, a different-contour trial might contrast Mary in a bassoon timbre with Twinkle in a bassoon timbre. No trials contained both contour and timbre differences.

Procedure. Children first received training trials on which highly-distinct melodies were used (rising vs. falling, high harp vs. low tuba—a 1.5 octave difference in pitch range), including four "same" trials and four "different" trials in each 8-trial training block. If they did not achieve at least 7/8 correct, the 8-trial training block repeated, up to 5 total training blocks. Children who never passed training were excluded from analysis. When criterion was achieved, the child continued to the test phase. The test phase presented all trials in a random order. Test stimuli included 8 melody-change trials (half familiar melodies, half scrambled), 8 timbre-change trials (half familiar melodies, half scrambled), and 16 same trials. An additional 8 trials presented the training stimuli (half same, half different) to assess continued task adherence. These trials were not analyzed.

Results

Prior to analysis, accuracy was converted to d-prime, a standard measure of change detection (McMillan & Creelman, 2005). Extreme values 0 and 1 were converted to $0+1/(2N) = .125$ and $1-1/(2N) = .875$, respectively, to avoid z-scores of ±infinity. Results appear in Figure 4. We then conducted an ANOVA on d-prime values, with independent variables Timbre Similarity (close, mid, far; between-subjects), Melody Familiarity (familiar, scrambled; within-subjects) and Trial Type (different melodies, different timbres; within-subjects).

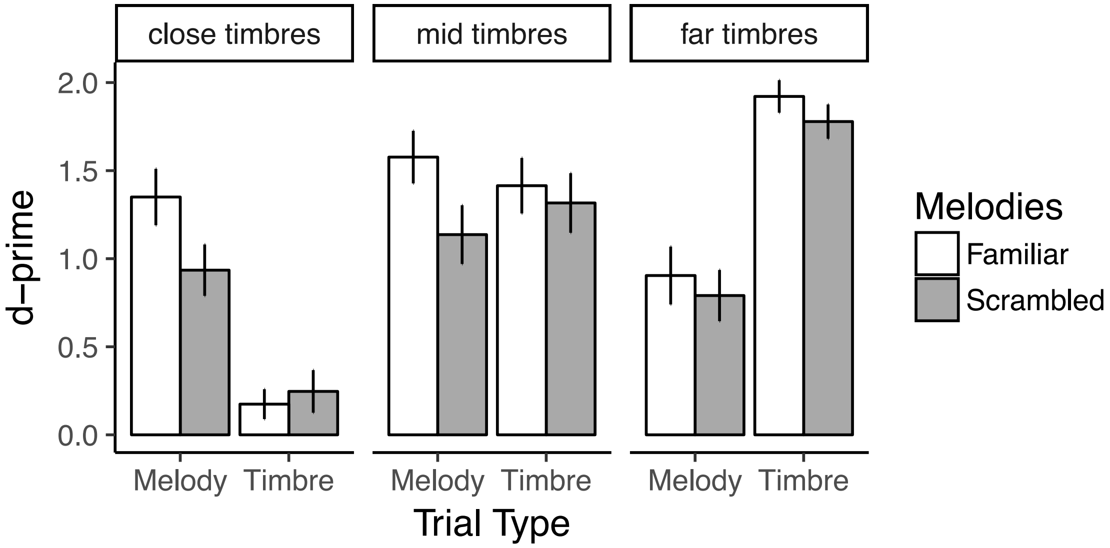

Figure 4. Experiment 2, discrimination accuracy, with standard errors. Plots created in R using ggplot 2 (Wickham, 2016).

There was an effect of Melody Familiarity (F(1,87)=15.85, p=.0001, η^2_P=.15), suggesting higher d-prime scores for familiar melodies. There was an also an effect of Timbre Similarity (F(2,87)=15.21, p<.0001, η^2_P=.26), suggesting that overall accuracy increased as the timbres became less similar. The Timbre Similarity main effect was qualified by an interaction of Timbre Similarity x Trial Type (F(2,87)=33.98, p<.0001, η^2_P=.44), which appeared to result from differences in the relative discriminability of timbres in the different conditions, such that melody discrimination exceeded timbre discrimination for close timbres, roughly matched it for moderately-similar timbres, and undershot timbre discrimination for far timbres. Finally, an interaction of Melody Familiarity x Trial Type (F(1,87)=5.88, p=.02, η^2_P=.06) suggests that the effect of melody familiarity was larger (.32 vs. .06) and significant (F(1,89)=18.78, p<.0001, η^2_P=.17) when children were discriminating melodies, but when they were discriminating timbres it was not significant (F(1,89)=0.71, p=.40, η^2_P=.01).

Discussion

Children appear to be better at distinguishing familiar melodies from each other than at distinguishing unfamiliar melodies from each other. This result held when melodies were played in a variety of timbres. This suggests that children have more robust perceptual representations of familiar melodies than of unfamiliar melodies. Interestingly, a large timbre difference still seems more salient than a difference between melodic contours, even when those melodic contours are familiar. This suggests that melody familiarity does not become more salient than substantial timbre differences, a salience pattern that has been previously reported in this age group (Creel, 2014b, 2016).

General Discussion

We originally asked whether children might perceptually represent music by storing entire melodies, rather than melodic properties, as on the specific learning hypothesis. If so, then familiar melodies should show perceptual processing advantages over matched unfamiliar melodies. This appears to be the case: Children are better at mapping familiar melodies to pictures than they are at mapping unfamiliar melodies to pictures, and they are better at judging perceptual distinctions between familiar melodies than between unfamiliar melodies. These findings are consistent with the specific learning hypothesis, and more generally with exemplar-style accounts of auditory pattern learning across domains.

The picture-association effect is particularly interesting in that two previous studies (Creel, 2014b, 2016) have found that children in this age range are unable to learn associations between novel melodic contours and novel pictures. Those findings suggest that pitch-contour differentiation in the context of a learning task is quite difficult. It seems that melody familiarity is sufficient to ease this task. Of course, one might reverse this observation to ask why children did not perform better at learning visual associations that were linked to familiar songs' lyrics. One answer may be that not all children were familiar with these songs, and only the ones who were familiar could learn the association. A different answer is that children may be disadvantaged here vs. recognizing familiar music in natural situations, because in natural listening situations they have access to additional distinguishing cues such as timbres and lyrics (see Vongpaisal, Trehub, & Schellenberg, 2009, for evidence that timbres and lyrics contribute to children's music recognition).

In sum, this research contributes to the literature on auditory perceptual development by suggesting that, as on the specific learning hypothesis, children are learning representations of specific auditory events, rather than simply accruing melodic properties.

Relationship to word learning

The research presented here also speaks to the literature on word learning by suggesting that the principles governing word learning are in fact broader perceptual-learning and/or associative learning principles. Recall two contrasting results in child word learning: while pictures with subtly-different novel word names (*bih, dih*) are not recognized by 14-month-olds (Stager & Werker, 1997), subtly-different familiar words (*ball, doll*) are easily recognized by the same age group (Fennell & Werker, 2003). We find that, in slightly older children, in an audiovisual association task that is similar to word learning, familiar melodies are more readily associated and recognized than unfamiliar melodies; and that in an immediate discrimination task, familiar melodies are more readily differentiated than unfamiliar melodies. The latter finding has not, to our knowledge, been demonstrated in young children's spoken word processing.

An open question is whether the role of familiarity is not only familiarity with the auditory form, but the existence of semantic associations. Clearly the familiar auditory form is necessary, in that semantic associations did not help in the presence of scrambled melodies. But are children using auditory familiarity alone, or some composite of auditory familiarity and semantic associations? In the picture-matching case (Experiment 1), the familiarity effect may depend on the naturalness of the picture mapping itself: picture associations with Twinkle Twinkle Little Star are learned best when the picture is a star, not when it is a novel cartoon character. This suggests that existing semantic associations have an influence on performance. Still, it is interesting that having semantic associations pointed out did not help children learning to associate the melodies with novel pictures. Had children simply associated the word "lamb" with one character and "star" with the other character, this presumably would have been an easy task for them, yet it was not. See Creel (2014a) on children's high learning accuracy for dissimilar verbal labels for the same cartoon characters.

The similarity judgments are easier to interpret as support for familiarity of auditory representations, as children could respond without activating meaning associations. Still, one might ask whether children performed better on familiar melodies in the discrimination task because they recoded real melodies into meaning-based representations, which they could not do for the unfamiliar melodies. This is certainly possible. However, it should be kept in mind that children were not very good at naming songs in pilot data even when prompted to do so (about 35% on the two selected songs), suggesting that lyrics recall, and therefore semantic associations, may not be automatic for children. Given that we did not tell children in Experiment 2 that they would be hearing familiar melodies, it is possible that many or most of them did not register that some songs were familiar.

To the extent that the semantic account of our results is valid, it may suggest a different parallel between word-form learning and auditory pattern learning. Specifically, the semantic account may indicate that external (semantic or lyrical) associations facilitate learning of auditory form, an account that has been proposed for word learning (e.g. Yeung & Werker, 2009). That is, perhaps association with distinct semantic networks sharpens or pattern-separates auditory representations of melodic properties.

Regardless of the exact nature of familiar-melody facilitation, though, it is clear that long-term familiarity of some sort—learning melodies and their associations over multiple days, weeks, or months—is necessary to see these effects. Creel (2014b, 2016) found little learning with multiple exposures in brief lab association experiments. Further, Creel's (2016) Experiment 1 presented children with extra exposure to the novel melodies prior to the association task, but even this extra exposure did not allow children to learn associations. Thus, extensive exposure but not brief exposure to specific auditory patterns may be necessary to see facilitated performance.

Limitations

A limitation to this study is the small number of melodies used. It remains possible that subtle differences between the familiar and unfamiliar melodies may have driven differences in effects. As discussed in the preliminary studies, the use of a large number of melodies is constrained by the number of melodies children can reliably recognize, and they had difficulty recognizing these ostensibly familiar melodies. While many child language experiments depend on a relatively small number of words, such as the foundational Stager and Werker (1997) study, it would be reassuring to see the current effect extended to other data points. One fruitful approach would be to test children in a single school setting where children's musical exposure repertoire is well-known, or, even better, in two different schools where children have learned two different sets of melodies, providing a fully crossed design. Such work should include a wider range of children's music, especially music without lyrics and without differing extramusical associations, to dissociate roles of semantic association and auditory familiarity.

A second limitation is that the current research does not make clear what properties children can use to organize representations of familiar melodies. We carefully selected melodies with identical timing properties, so that melodic patterns were the focus of inquiry. Of course melodic patterns include not just contour, but exact pitch distances, scale degrees included, pitch range, and tonality (major/minor, which did not vary here). What we have shown here is that well-known melodies themselves, not just schematic melodic features, may be an organizing principle. Future work should examine additional factors in children's musical representations.

Open data

Data and R code for analysis are available at the following DOI 10.17605/OSF.IO/WH3FY

References

Bates, D., Maechler, M., Bolker, B., & Walker, S. (2015). Fitting linear mixed-effects models using lme4. *Journal of Statistical Software, 67*(1), 1–48.

Boersma, P., & Weenink, D. (2014). Praat: Doing phonetics by computer [Computer program]. Version 5.4.01. Available at: http://www.praat.org/. Accessed Nov. 9, 2014.

Corrigall, K. A., & Trainor, L. J. (2010). Musical enculturation in preschool children: acquisition of key and harmonic knowledge. *Music Perception*, *28*(2), 195–200.

Creel, S. C. (2014a). Preschoolers' flexible use of talker information during word learning. *Journal of Memory and Language, 73*, 81–98. http://doi.org/http://dx.doi.org/10.1016/j.jml.2014.03.001

Creel, S. C. (2014b). Tipping the scales: Auditory cue weighting changes over development. *Journal of Experimental Psychology: Human Perception and Performance*, *40*(3), 1146–1160. doi:10.1037/a0036057

Creel, S. C. (2014c). Impossible to _gnore: Word-Form Inconsistency Slows Preschool Children's Word-Learning. *Language Learning and Development, 10*(1), 68–95. http://doi.org/10.1080/15475441.2013.803871

Creel, S. C. (2016). Ups and downs in auditory development: Preschoolers' sensitivity to pitch contour and timbre. *Cognitive Science*, *14*, 373–303. doi:10.1111/cogs.12237

Creel, S. C., & Jimenez, S. R. (2012). Differences in talker recognition by preschoolers and adults. *Journal of Experimental Child Psychology, 113*, 487–509.

Diehl, R. L., & Walsh, M. A. (1989). An auditory basis for the stimulus-length effect in the perception of stops and glides. *Journal of the Acoustical Society of America, 85*(5), 2154–2164. http://doi.org/10.1121/1.397864

Doherty, M. J. (2004). Children's difficulty in learning homonyms. *Journal of Child Language*, *31*, 203–214. http://doi.org/10.1017/S030500090300583X

Dowling, W. J. (1978). Scale and contour: two components of a theory of memory for melodies. *Psychological Review*, *85*(4), 341–354.

Fennell, C. T., & Werker, J. F. (2003). Early word learners' ability to access phonetic detail in well-known words. *Language and Speech, 46*(Pt 2-3), 245–64. http://doi.org/10.1177/00238309030460020901

Fragoulis, D., Papaodysseus, C., Exarhos, M., & Roussopoulos, G. (2006). Automated classification of piano-guitar notes. *IEEE Transactions on Audio, Speech, and Language Processing, 14*(3), 1040–1050. http://doi.org/10.1109/TSA.2005.857571

Goldinger, S. D. (1998). Echoes of echoes?: An episodic theory of lexical access. *Psychological Review*, *105*(2), 251–279.

Hannon, E. E., & Trehub, S. E. (2005). Tuning in to musical rhythms: Infants learn more readily than adults. *Proceedings of the National Academy of Sciences*, *102*(35), 12639–12643.

Hannon, E. E., & Trehub, S. E. (2005). Metrical categories in infancy and adulthood. *Psychological Science*, *16*(1), 48–55.

Hay, J. S. F., & Diehl, R. L. (2007). Perception of rhythmic grouping: Testing the iambic/trochaic law. *Perception and Psychophysics, 69*(1), 113–122. http://doi.org/10.3758/BF03194458

Iverson, P., & Krumhansl, C. L. (1993). Isolating the dynamic attributes of musical timbre. *Journal of the Acoustical Society of America*, *94*(5), 2595–603. Retrieved from http://www.ncbi.nlm.nih.gov/pubmed/8270737

Kluender, K. R., Diehl, R. L., & Killeen, P. R. (1987). Japanese quail can learn phonetic categories. *Science, 237*(4819), 1195–1197.

Kluender, K. R., Diehl, R. L., & Wright, B. A. (1988). Vowel-length differences before voiced and voiceless consonants: An auditory explanation. *Journal of Phonetics*, *16*, 153-169.

Kuhl, P. K., Williams, K. A., Lacerda, F., Stevens, K. N., & Lindblom, B. (1992). Linguistic experience alters phonetic perception in infants by 6 months of age. *Science*, *255*(5044), 606–608.

Lynch, M. P., Eilers, R. E., Oller, D. K., Urbano, R. C., & Oiler, D. K. (1990). Innateness, experience, and music perception. *Psychological Science, 1*(4), 272–276.

MacMillan, N.A., & Creelman, C.D. (2005). *Detection theory: A user's guide*. Mahwah, NJ: Taylor & Francis.

Pajak, B., Creel, S. C., & Levy, R. (2016). Difficulty in learning similar-sounding words: a developmental stage or a general property of learning? Journal of Experimental Psychology: Learning, Memory, and Cognition, 42(9), 1377–1399. http://doi.org/10.1037/xlm0000247

R Core Team (2014). *R: A language and environment for statistical computing*. Vienna, Austria: R Foundation for Statistical Computing. URL: http://www.R-project.org.

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.

Storkel, H. L., & Maekawa, J. (2005). A comparison of homonym and novel word learning: The role of phonotactic probability and word frequency. *Journal of Child Language*, *32*(4), 827–853. http://doi.org/10.1017/S0305000905007099

Swingley, D., & Aslin, R. N. (2002). Lexical neighborhoods and the word-form representations of 14-month-olds. *Psychological Science*, *13*(5), 480–484. doi:10.1111/1467-9280.00485

Trainor, L. J., & Trehub, S. E. (1994). Key membership and implied harmony in Western tonal music: developmental perspectives. *Perception & Psychophysics*, *56*(2), 125–32. Retrieved from http://www.ncbi.nlm.nih.gov/pubmed/7971113

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

Vongpaisal, T., Trehub, S. E., & Schellenberg, E. G. (2009). Identification of TV tunes by children with cochlear implants. *Music Perception*, *27*(1), 17–24.

Werker, J. F., Fennell, C. T., Corcoran, K. M., & Stager, C. L. (2002). Infants' ability to learn phonetically similar words: Effects of age and vocabulary size. *Infancy*, *3*(1), 1–30. doi:10.1207/15250000252828226

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*(1), 49–63. http://doi.org/10.1016/S0163-6383(84)80022-3

Wickham, H. (2016). *ggplot2: Elegant Graphics for Data Analysis*. Springer-Verlag New York.

Yeung, H. H., & Werker, J. F. (2009). Learning words' sounds before learning how words sound: 9-month-olds use distinct objects as cues to categorize speech information. *Cognition*, *113*(2), 234–43. http://doi.org/10.1016/j.cognition.2009.08.010