



PAPER

Everyday music in infancy

Jennifer K. Mendoza | Caitlin M. Fausey

Department of Psychology, University of Oregon, Eugene, Oregon, USA

CorrespondenceCaitlin M. Fausey, Department of Psychology, 1227 University of Oregon, Eugene, OR, 97403, USA.
Email: fausey@uoregon.edu**Abstract**

Infants enculturate to their soundscape over the first year of life, yet theories of how they do so rarely make contact with details about the sounds available in everyday life. Here, we report on properties of a ubiquitous early ecology in which foundational skills get built: music. We captured daylong recordings from 35 infants ages 6–12 months at home and fully double-coded 467 h of everyday sounds for music and its features, tunes, and voices. Analyses of this first-of-its-kind corpus revealed two distributional properties of infants' everyday musical ecology. First, infants encountered vocal music in over half, and instrumental in over three-quarters, of everyday music. Live sources generated one-third, and recorded sources three-quarters, of everyday music. Second, infants did not encounter each individual tune and voice in their day equally often. Instead, the most available identity cumulated to many more seconds of the day than would be expected under a uniform distribution. These properties of everyday music in human infancy are different from what is discoverable in environments highly constrained by context (e.g., laboratories) and time (e.g., minutes rather than hours). Together with recent insights about the everyday motor, language, and visual ecologies of infancy, these findings reinforce an emerging priority to build theories of development that address the opportunities and challenges of real input encountered by real learners.

KEYWORDS

enculturation, everyday ecologies, infancy, input, LENA, music

1 | INTRODUCTION

Humans begin to enculturate to their musical soundscape early in infancy. Sensitivities to rhythm and scale structures specific to an infant's culture are evident around their first birthday (e.g., Hannon & Trehub, 2005a, 2005b; Lynch & Eilers, 1992; Soley & Hannon, 2010) and the encountered musical notes and rhythms continue to shape musical sensitivities from infancy through adolescence (e.g., Hannon & Trainor, 2007; Trainor & Corrigan, 2010). How could caregivers singing, instruments riffing, siblings clapping, voices harmonizing, and tunes repeating shape this enculturation? We do not currently understand the role of sensory history in emerging sensitivities to surrounding

sounds in part because we have no model of the sensory history. In order to have a well-specified theory of enculturation, we need to know what features, voices, and tunes are available in the musical ecology.

Models of the ecologies in which organisms build skills are required for theories of development (Gottlieb, 1991; West & King, 1987) because sensory history shapes emerging neural circuitry and skills (e.g., Aslin, 2017; Hensch, 2005; Scott et al., 2007; Werker & Hensch, 2015). Useful models of the sensory history are those that quantify the range and distribution of features encountered in real life. Recent discoveries about human motor, linguistic, and visual ecologies have taught us that such everyday distributions are not apparent in canonical researcher-constrained activities (e.g., Adolph et al., 2018; Roy

This is an open access article under the terms of the [Creative Commons Attribution-NonCommercial](https://creativecommons.org/licenses/by-nc/4.0/) License, which permits use, distribution and reproduction in any medium, provided the original work is properly cited and is not used for commercial purposes.

© 2021 The Authors. *Developmental Science* published by John Wiley & Sons Ltd



et al., 2015; Smith et al., 2018). For example, simplified walking tasks do not reveal frequent omnidirectional steps taken by freely walking infants (Lee et al., 2018). Short language tasks similarly miss properties of everyday language use like talking interleaved with silences (Tamis-LeMonda et al., 2017). Likewise, visual objects in most behavioral and machine learning training regimes are uniformly available, unlike object frequency distributions that are non-uniform and changing over time in everyday life (Clerkin et al., 2017; Fausey et al., 2016; Smith & Slone, 2017). To the extent that researcher-constrained activities contrast with the time, space, and/or feature distributions of everyday life, they are ill-suited to answer questions about the sensory history that shapes development. What is the musical ecology of everyday infancy?

1.1 | Infants' musical experiences, preferences, and skills

Infants encounter music during informal everyday activities (Trehub et al., 1997), with two-thirds of parents reporting that they had never attended a formal baby music class (Fancourt & Perkins, 2018). According to parent report, infants encounter music on a daily basis (Custodero & Johnson-Green, 2003). Caregivers report primarily singing to their infants (Ilari, 2005) and also describe infants listening to recorded music from TV shows, radios, and toys (Young, 2008). Recent advances in caregiver-report measures will likely reveal even more about infants' everyday musical activities (Politimou et al., 2018). Though caregiver report provides some insights into the everyday music of infancy, such estimates yield a sparse model of sensory history. The features that should organize more detailed models are readily suggested by studies documenting infants' musical preferences and skills.

We know that infants learn from vocal, instrumental, live, and recorded music in researcher-constrained tasks. For example, infants track tempo and pitch height in vocal music (e.g., Conrad et al., 2011; Volkova et al., 2006). Infants differentiate novel from familiar melodies that are instrumental (e.g., Trainor et al., 2004). They also recognize and remember melodies presented by both live and recorded sources (e.g., Mehr et al., 2016). Infants prefer music with only a voice compared to music that is both vocal and instrumental (e.g., Ilari & Sundara, 2009) and children have stronger memory for vocal melodies than for instrumental melodies (e.g., Weiss et al., 2015). Models of everyday music in infancy should therefore quantify opportunities for infants to learn from vocal, instrumental, live, and recorded music as well as rates of music instantiated with one (e.g., exclusively vocal) or more (e.g., vocal and instrumental) of these features.

Evidence from researcher-constrained activities also suggests that infants are sensitive to repetition in the music that they encounter. For example, infants' physiological arousal decreases when mothers repeatedly sing a soothing tune (Cirelli et al., 2019). Infants remember tunes that they have heard repeatedly, across delays as long as 14 days (e.g., Saffran et al., 2000). A repeatedly encountered tune often attracts further listening (e.g., Ilari & Polka, 2006; Plantinga & Trainor, 2009) and sometimes is sufficiently familiar to drive infants to sample another tune (e.g., Plantinga & Trainor, 2005; Saffran et al.,

Research Highlights

- We captured daylong recordings from 35 infants ages 6–12 months at home and identified everyday music and its features, tunes, and voices.
- Most instances of this everyday music were recorded and instrumental (e.g., toys) or live and vocal (e.g., caregiver singing).
- One musical tune or voice was much more available than many others; infants did not encounter each individual identity equally often in their day.
- Quantifying everyday music provides a foundation for future studies of musical enculturation grounded in sensory history.

2000). Infants are also more likely to help someone who had sung an especially familiar tune (Cirelli & Trehub, 2018). There is little direct evidence about infants' sensitivity to variation in tunes and voices. However, infants track features like timbre (e.g., Trainor et al., 2011) and meter (e.g., Hannon & Trehub, 2005b) across a variable set of tunes. Models of everyday music in infancy should therefore quantify opportunities to build skills across repeating and varying musical identities (i.e., individual tunes and voices).

1.2 | Quantifying everyday music

How often do infants encounter vocal, instrumental, live, and recorded music in everyday life? How available is each tune and voice identity in this everyday music? We do not yet know the musical ecology of everyday infancy. Quantifying this soundscape will reveal properties of infants' sensory histories that are available to shape infants' learning. These properties should be central to any theory of musical enculturation. One related domain in which we have learned a great deal from quantifying everyday distributions is object name learning. We now know that frequency matters, such that infants learn the most frequently encountered nouns earliest (Goodman et al., 2008). Range, as indexed by encountering many different words as well as words in varied contexts, also facilitates learning (Hills et al., 2010; Pan et al., 2005). The distributional shape of encountered objects also matters, such that infants learn the names of the most visually available objects earliest and may benefit from encountering a few objects a lot together with many others less frequently (Clerkin et al., 2017). We therefore aim to quantify the frequency, range, and distributional shape of everyday music in infancy as a foundational step toward revising theories of musical enculturation to include sensory history.

To do so, we captured infants' everyday soundscapes by audio recording full days at home. Initial insights from case studies of audio recordings at home suggest that infants will encounter an uneven soundscape of musical features. For example, across four day-long observations from two families, roughly 20–30 min of music per day

was live while 60–510 min of music per day was recorded (Costa-Giomi, 2016; Costa-Giomi & Benetti, 2017; Costa-Giomi & Sun, 2016). Here, we reveal the separate and joint availability of vocal, instrumental, live, and/or recorded music in infants' days. Within this feature space, infants will encounter individual tunes and voices. Evidence from other domains of everyday infancy suggests that not all identities will be equally available. For example, though infants encounter roughly eight unique face identities per four hours, proportionally nearly all of those face instances are from three or fewer people (Jayaraman et al., 2015). Similarly, among the many objects that infants encounter in mealtime activities, a small set are especially pervasive (Clerkin et al., 2017). The fact that caregivers can identify tunes that they sing repeatedly to their infants (Bergeson & Trehub, 2002) suggests that everyday distributions of musical identities may be similarly non-uniform. Here, we discover the extent to which infants encounter the same versus unique tunes (voices) per day.

1.3 | A first-of-its-kind snapshot of musical sensory history

We sampled at-home soundscapes of infants who were between 6- and 12-months-old in order to capture the sensory history available to shape musical enculturation. We sampled raw audio (Ford et al., 2008) in order to avoid estimation errors of self-report such as the discrepancy between a caregiver reporting singing "all the time" and a day-long audio recording of her family at home revealing less than 2 min of caregivers singing (Costa-Giomi & Benetti, 2017; Costa-Giomi & Sun, 2016). We sampled full days in order to discover potential repetition and variation of musical tune and voice identities. In order to advance theories of enculturation grounded in sensory history, we quantified the availability of live, recorded, vocal, and instrumental music as well as the tune and voice identities within each recording.

A complete understanding of the soundscapes of human infancy will require aggregating snapshots across many places and times (Hruschka et al., 2018; Nielsen et al., 2017). We suggest that fruitful aggregation will arise from snapshots that share three key properties of the present research: 1) sensory history is sampled without researchers present so that it is not artificially distorted, 2) many hours per day are sampled so that patterns of repetition and variation are discoverable, and 3) distributions are quantified so that future efforts to understand causal consequences of sensory history can manipulate both the frequency and diversity of instances in ways that infants encounter as they build knowledge.

2 | METHOD

2.1 | Ethics and open science statement

The University of Oregon Institutional Review Board approved this research protocol. Caregivers provided informed consent for their family's participation.

Most parents consented to share their daylong recordings with the research community and these .wav files are available on HomeBank (Fausey & Mendoza, 2018). Study materials, behavioral coding manuals, numerical data, and analysis code are available on Open Science Framework (Mendoza & Fausey, 2019, henceforth "OSF").

2.2 | Participants

One daylong recording from each of 35 infants between the ages of 6 and 12 months ($M = 38.78$ weeks, $SD = 6.66$ weeks) was coded and analyzed. Families' race, income, and education were distributed as in the local community (U.S. Census Bureau, 2016).

The set of recordings was designed to meet the following criteria: infant ages, assigned sexes, and days of the week were roughly evenly represented, and each recording captured at least 10 h of everyday life. Most families recorded on their pre-assigned day of the week; we retained recordings from families who deviated due to family circumstances ($N = 2$). Short recordings were not analyzed ($N = 6$) and another family was recruited. Families also contributed to a larger project recording three days per week; which day was coded for the music corpus was determined before researchers listened to any of the family's recordings. Families received \$50 and a children's book.

2.3 | Materials

In order to capture each infant's full day of sounds, we used the digital language processor (DLP) from the Language Environment Analysis system (Ford et al., 2008) that records up to 16 h of audio. Each family received a DLP, infant vest, and diary log for their recording day.

Caregivers completed questionnaires after their recording day (MBCDI, Fenson et al., 1993, and custom queries about their family; OSF).

2.4 | Procedure

A researcher met with caregivers before their recording day in order to provide study materials and instructions. Caregivers were told that the study was about the mix of sounds infants hear in their natural environments (e.g., people talking, radios playing, dogs barking, refrigerators running, etc.). Caregivers were not told that music would be an analytic target.

After caregivers learned how to turn on the DLP and place it inside the infant vest, they also learned three goals for their recording day: turn on the DLP when their infant first woke up in the morning and leave it on until their infant went to sleep at night, remove the vest during naps and baths, but leave the DLP turned on and nearby, and remove the vest before traveling outside the home due to Oregon state laws about audio recording in public. Caregivers used the diary log to note any times when these situations happened, as well as any periods of time that they wanted researchers to delete from the recording due to private content.



The researcher called caregivers on their recording day to provide a chance to discuss any questions and also met with caregivers after the recording day to collect materials, administer questionnaires, and debrief.

2.5 | Data pre-processing

Each recorded day yielded a .wav file of up to 16 h duration. We edited .wav files as necessary to maintain privacy and preserve clock time. We replaced original sounds with silence during any times that caregivers logged as private or outside the home and we inserted silence between any times that the DLP was turned off and back on again. We also identified long stretches of silence as unavailable for coding (OSF; Bergelson & Aslin, 2017).

2.6 | Coding music

Automatic identification of music in recordings from everyday life is not yet possible. Thus, human coders identified bouts of music and tagged their features, voices, and tunes. In total, the efforts of 38 coders (approximately 6,400 person hours) yielded a first-of-its-kind corpus of everyday music in infancy.

2.7 | Identifying music bouts

We defined “music” as live singing (e.g., caregivers, siblings) and/or instrument playing (e.g., piano, guitar), recorded singing and/or instrument playing (e.g., radio, toys), and pitched, rhythmic, repetitive patterns that were vocally produced (e.g., humming, whistling, “vocal play”) and/or instrumental (e.g., clapping, drumming). Sounds that were produced by the focal infant (determined by contextual cues), speech (including infant-directed speech and routinized speech like book reading), infant babbling and/or an imitation of infant babbling, sound effects (e.g., “beep beep”), sounds of non-musical household objects (e.g., computer keyboard, microwave), and sounds produced by a non-human animal (e.g., birdsong) were not coded as music.

Coders used ELAN Linguistic Annotator (Version 4.9.4; Wittenburg et al., 2006) to listen continuously to the daylong audio recording. Upon hearing a musical sound, coders marked the onset and offset of the music bout. We defined a bout as the uninterrupted, continuous presence of music. Music bouts were determined independently from the musical content present. For example, a bout could include one, two, or more tunes; likewise, a single tune could be split across multiple (interrupted) bouts. Bouts ended when the source of the music stopped producing musical sounds or the musical sounds became too faint or too obscured to be perceived by the coder.

Coders were trained with a coding manual (OSF) and also reviewed the manual at the start of every coding session. Two independent coders identified music in each recording.

2.8 | Identifying features, voices, and tunes

Coders identified the features, voices, and tunes of music bouts in multiple passes.

First, coders judged whether any music in each bout was live and/or recorded. “Live” music bouts contained a musical sound produced by a human who was clearly present in-person in the infant’s environment (e.g., human voice, human rhythm like clapping, live instrument playing). “Recorded” music bouts contained a musical sound produced by an electronic source (e.g., TV, Pandora, toy).

Next, in a separate coding pass, coders judged whether each music bout contained any vocal and/or instrumental music. “Vocal” music bouts contained a musical sound that was produced by a live or recorded voice (e.g., adult, non-focal child, or recorded character singing, humming, whistling, vocal play). “Instrumental” music bouts contained a musical sound produced by a live or recorded instrument (e.g., piano, guitar, non-vocal musical sounds from toys, and non-vocal musical sounds like clapping).

Next, coders identified the specific voice(s) that produced music in each “vocal” bout (e.g., Mom, Grandparent, Taylor Swift, Daniel Tiger). If coders did not know the specific voice, they searched the Internet using available cues and knowledge. They did not use software like Shazam that required direct access to the recording, nor did they discuss it with colleagues or friends, in order to maintain confidentiality of each family’s recording. If coders could not determine the specific voice identity, then they created a distinct label (e.g., Female voice 1, Squeaky cartoon voice 2). As they proceeded through the vocal bouts of a recording, coders judged whether the current voice was the same as or different from all previously coded voices in the recording. If it was the same, then coders listed the same specific identity as when the voice occurred previously (e.g., Mom and Mom). If it was different, then coders listed a unique identity (e.g., Mom and Female voice 2). Critically, if coders encountered repeated instances of the same voice across vocal bouts within a recording, then they listed exactly the same identity for each instance of the same voice.

Finally, coders identified the tune(s) that occurred in each music bout. Every music bout had at least one tune. Coders listed the known title if discoverable (e.g., Itsy Bitsy Spider, Shake It Off) or they created a short, descriptive title for the tune (e.g., Everybody loves potatoes, Short Whistle 4). As with voices, coders judged whether the current tune was the same as or different from all previously coded tunes in the recording. If it was the same, then coders listed the same specific title as when the tune occurred previously (e.g., Itsy Bitsy Spider and Itsy Bitsy Spider). If it was different, then coders listed a unique specific identity (e.g., Itsy Bitsy Spider and Fast pop song 3). Critically, if coders encountered repeated instances of the same tune across music bouts within a recording, then they listed exactly the same title for each instance of the same tune.

Coders listened to the music bouts previously tagged in ELAN and entered their features, voices, and tunes coding into one Excel file per coding pass. Pilot coding revealed that identifying musical content was taxing and also that many music bouts had a single feature, voice, and tune. For these reasons, we focused on identifying content, rather than



additionally parsing bout-internal timing for any bouts with multiple features, voices, and/or tunes. Coders were trained with a one-time media review, a manual, and also reviewed the manual at the start of every coding session (OSF). Two independent coders identified the features, voices, and tunes of each music bout.

2.9 | Training-to-criterion and assessing reliability

All coders successfully reached criterion agreement with an expert coder before coding primary data.

Six daylong recordings (collected for training purposes and not analyzed) served only as training files. The expert coder identified music bouts in three recordings and the features, voices, and tunes in the music bouts of the other three recordings. These training files contained a range of voices and tunes in all combinations of features, including bouts with multiple voices and/or multiple tunes. All coders first coded at least one training file for music bouts and at least one training file for features, voices, and tunes. If coders failed to reach criterion on their first training file, they received feedback and could code up to two additional training recordings.

For music bouts, coders passed training if the number of seconds coded as music per each minute of the training file correlated at least $r = .90$ with the expert codes. For the features “live”, “recorded”, “vocal”, and “instrumental”, coders passed training if the proportion agreement with the expert codes across all music bouts was at least $.90$. For voices and tunes, the analytic target was distributional structure across a daylong recording. Contingency between the trainee’s and expert’s coding was therefore assessed. For example, the expert coder could have labeled the voices in three separate music bouts as “Friend”, “Friend”, and “Friend”, while the trainee could have labeled the voices in those same three music bouts as “Neighbor”, “Neighbor”, “Neighbor”. If we were to determine reliability based on whether the identity labels matched, then we would find no agreement for these three music bouts. However, the expert and trainee did show some agreement – they each coded these three music bouts with only a single unique voice identity label, internal to their own coding. Therefore, the *distributional structure* of their coding was reliable. In contrast, if the trainee had labeled the voices in these three music bouts as “Neighbor”, “Pop Singer”, and “Sibling”, then neither the labels nor the structure of their coding would agree with that of the expert. Assessing contingency for voices and tunes coding allowed us to evaluate reliability based on distributional structure, not identity labels, between the trainee’s and the expert’s coding. We first screened voice and tune identity labels for consistency in spelling, capitalization, spacing, and punctuation and then assessed contingency between the trainee and the expert labels. We assessed contingency using Tschuprow’s T , which ranges from 0 to 1, reaches 1 only in the case of a square table, and otherwise permits rectangular tables (e.g., if the expert and trainee identified a different total number of unique voices throughout the recording). Coders passed training when Tschuprow’s T was at least $.90$ between their coding and the expert coding, separately for coding both voice and tune identities.

Primary data were coded by two independent coders who had passed training. Reliability was assessed using the same measures as for training-to-criterion.

2.10 | Music and its features, voices, and tunes across the day’s seconds

Each recording is represented as a timeseries of seconds. Linked timeseries indicate when the DLP was recording, when the original .wav file was edited to replace private or outside-the-home episodes with silence, when the recording was coded for music, and the results of each music coding pass. Music bout onsets and offsets were converted from ELAN coding into this format by rounding onsets down and offsets up to the nearest second. When two ELAN-coded bouts were less than one second apart ($.13$ of ELAN bouts), they were merged into one bout. Each second within a bout inherited its features, voices, and tunes.

3 | RESULTS

3.1 | Capturing everyday sounds

We captured 466.76 h of the everyday sounds of infancy. The corpus consists of 35 recordings each capturing an infant’s full day (*Median* = 13.13 h per day; *Interquartile range (IQR)* = 4.32 h per day). After pre-processing (Method), coders listened to 269.69 h in order to identify music and its features, voices, and tunes (*Median* = 8.07 coded h per day, *IQR* = 2.89 coded h per day).

3.2 | Reliably identifying music and its features, voices, and tunes

Two coders independently identified the music bouts in each recording. The number of seconds coded as music per each minute by the coders of each recording was highly correlated (*Median* $r = .93$, *IQR* $r = .11$; see Figure 1 for one example daylong timeseries). Because inter-rater reliability was high, we randomly selected one coder’s music bouts per recording to further code for features, voices, and tunes.

Two coders independently coded the features, voices, and tunes that occurred in music for each recording. Rarely ($n = 18$ bouts; 169 s), these coders did not discern a musical sound in a previously identified music bout. These bouts were not coded or analyzed for features, voices, and tunes. The proportion agreement between coders for each recording was high for all features: live music (*Median* = $.98$, *IQR* = $.03$), recorded music (*Median* = $.99$, *IQR* = $.02$), vocal music (*Median* = $.98$, *IQR* = $.03$), and instrumental music (*Median* = $.99$, *IQR* = $.04$). We assessed the reliability of identifying specific voices and tunes using Tschuprow’s T (as when training coders to criterion, see Methods). Inter-rater reliability for each recording was high for both tunes (*Median* = $.90$, *IQR* = $.06$) and voices (*Median* = $.94$, *IQR* = $.11$; see Figure 1 for one example rectangular contingency table). Because

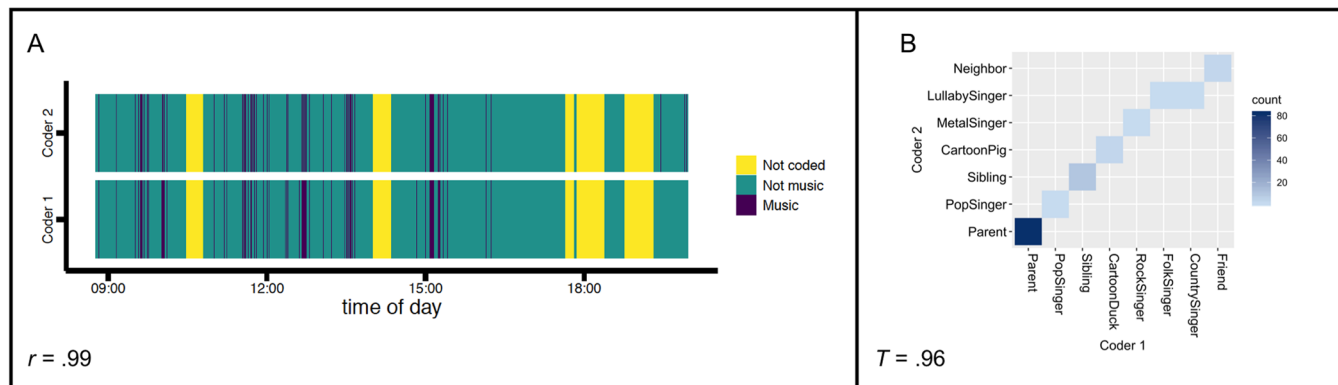


FIGURE 1 Example of high inter-rater reliability for identifying music and the distributional structure of its voices in one daylong recording. (A) Music bouts: Coder 1 and Coder 2 identified music bouts (purple) throughout the recording. (B) Voices: Coder 1 and Coder 2 identified the specific voices in this recording's vocal bouts (voice identities printed here are fictional, in order to maintain family privacy). Blue tiles show the intersection of Coder 1's and Coder 2's identity for each bout; darker tiles indicate a larger number of bouts. This illustration of contingency shows reliable distributional structure (e.g., Coder 1's 'Friend' and Coder 2's 'Neighbor' uniquely identified this voice) with one minor deviation from perfect agreement (e.g., Coder 1 distinguished "FolkSinger" from "CountrySinger" while Coder 2 identified both voices as "LullabySinger").

inter-rater reliability for coding features, voices, and tunes was high, we randomly selected one coder's features, voices, and tunes per recording for analysis. Prior to analysis, all recordings were screened by an independent coder for any internal inconsistencies (e.g., a bout coded as "vocal" with no voice identity listed); as expected, these were rare (.002 live bouts; .002 recorded bouts; .010 vocal bouts; .011 instrumental bouts) and easily resolved when the independent coder listened to each bout.

3.3 | Everyday music in infancy: Overview

We present new discoveries about the everyday music available to infants in the second half of their first postnatal year, based on 35 daylong recordings sampled from this developmental period. We first describe rates of live, recorded, vocal, and instrumental features in everyday music. We then detail distributions of individual tune and voice identities within daily music.

3.4 | The feature space of everyday music in infancy

We identified 4,798 bouts of music in this corpus (*Median* = 127 bouts per recording, *IQR* = 120 bouts per recording). The source of music was vocal in over half, and instrumental in over three-quarters, of infants' everyday music bouts. Music was generated by live sources in approximately one-third of the bouts and by recorded sources in roughly three-quarters of the bouts (Figure 2).

The feature space within individual recordings is consistent with the pattern shown in Figure 2. Recorded instrumental music and/or live vocal music are in the two densest cells of the feature space for 34 of the individual recordings, with the remaining recording densest in combinations of vocal, instrumental, recorded, and live music (OSF).

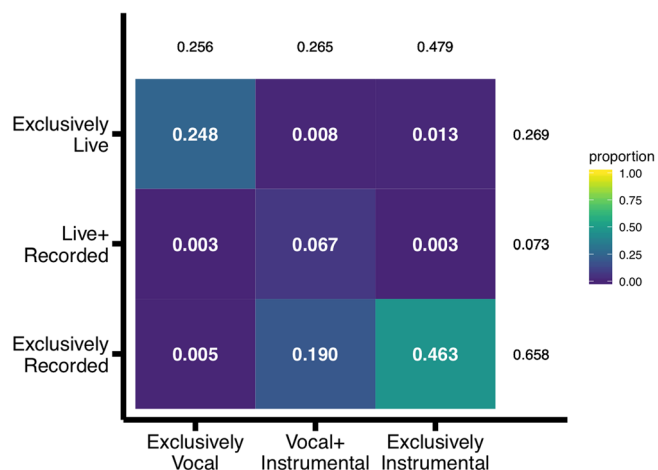
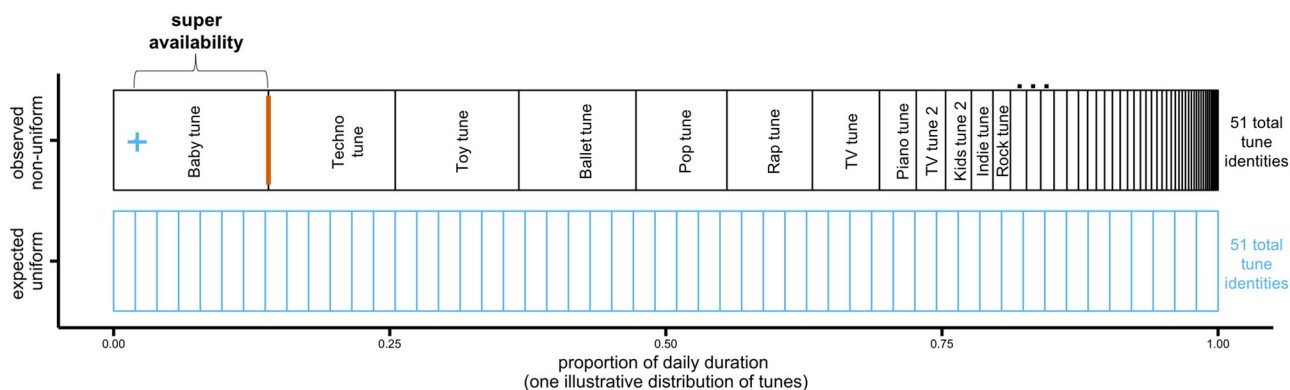


FIGURE 2 The feature space of everyday music in infancy. Each cell shows the proportion of bouts containing music with the features of its row by column intersection. Lighter colors indicate denser parts of the feature space.

Figure 2 depicts bout-space. Bouts cumulated to a total of 42.01 h of music in this corpus (*Median* = 3,311 s per recording; *IQR* = 3,858.50 s per recording). Converting the bout-space to duration-space is exact for bouts with one feature (e.g., "exclusively live"; all of the seconds in the bout inherit the feature). Converting the bout-space to duration-space is less exact for bouts with multiple features. For example, imagine a bout coded as "vocal" and "instrumental" because it had both a caregiver singing and another caregiver playing guitar. One possibility is that both music sources persisted for the whole bout – the caregiver singing along with the guitar. Another possibility is that the voice and guitar each persisted for half of the bout – the caregiver singing solo followed by the other caregiver playing guitar solo. We discovered that the bulk of music bouts contained only one or the other value of each dimension (.93 exclusively live or exclusively recorded; .73 exclusively vocal or exclusively instrumental). Duration-space figures in which all

TABLE 1 Tunes and voices per day in the everyday music in infancy

	Total music seconds per day			Number of unique identities per day		
	Median	IQR	Min-Max	Median	IQR	Min-Max
Tunes	2404	3369.5	361-9309	51	36.5	14-213
Voices	422	773	72-2001	3	1	1-5

**FIGURE 3** One illustrative distribution of tunes showing an observed daily distribution (top row, black) that is not consistent with a uniform identity space (bottom row, blue). The blue plus sign shows the proportion for one tune identity expected from uniformly distributed identities and the thick orange line shows the proportion of the most available observed tune identity (“Baby tune”). See text for super availability computation.

seconds inherited all of its bouts’ features (potentially overestimating certain features’ durations, but minimally so) are presented as supplemental information on OSF.

3.5 | The identity space of everyday music in infancy

Specific tunes and voices (i.e., musical identities) were detected within each recording and so our unit of analysis is the daily distribution. We report daily distributions of (a) Tunes: bouts with a single tune (99,157 total seconds; .84 of all bouts) and (b) Voices: bouts that were exclusively vocal with a single voice (21,735 total seconds; .24 of all bouts). We report on bouts in which one tune, voice, and/or feature was identified, yielding certainty that each tune (voice) persisted for the whole bout. Distributional analyses that also include bouts with multiple tunes, voices, and/or features show similar patterns (OSF).

We test whether observed daily distributions are consistent with a uniform identity space – in which each individual tune (voice) is equally available. Table 1 shows the cumulated seconds and number of unique identities per day in the tune identity space and in the voice identity space. We use these observations to generate expected distributions of uniformly available tune (voice) identities.

Figure 3 depicts an illustrative example of how we compare observed and expected distributions. In Figure 3, we depict a fictional day with 51 tune identities. 51 tunes divided uniformly would mean that each tune accounts for .02 of the distribution ($1/51 = .0196$). In contrast to this expected uniform, we depict an observed distribution

with one tune (“Baby tune”) accounting for .14 of the daily distribution. The extent to which this most available observed identity exceeds the uniform expected per-identity is the “super availability”. Thus, in Figure 3, the super availability is $.12 : .02$ (expected per tune) + $.12$ (super availability) = $.14$ (observed most available tune). Another way to think about the difference between these distributions is that the most available tune in this fictional day is 7 times more available than would be expected if each tune had occurred equally ($.14/.02 = 7$).

We test whether observed daily distributions are consistent with a uniform identity space at two different scales. We first examine this at the recording scale, comparing the observed daily distribution of tune (voice) identities in each individual recording to its yoked uniform distribution. This allows us to detect the extent to which observed distributional patterns are a general property of individual infants’ daily distributions of tunes and voices. We then examine this at the corpus scale, aggregating the observed identity durations (e.g., a 50-s tune, a 10-s tune, a 22-s tune, and so on) from all 35 recordings and aggregating the expected identity durations from all 35 recordings. This allows us to discover new insights about likely tune (voice) identity durations in infants’ – not just one infant’s – everyday musical ecologies.

3.6 | Recording scale: Tune and voice identities are not uniformly available

We discovered that identities cumulated to more extreme longer and shorter durations per day than would be expected under the assumption that every identity was available for the same duration per day.

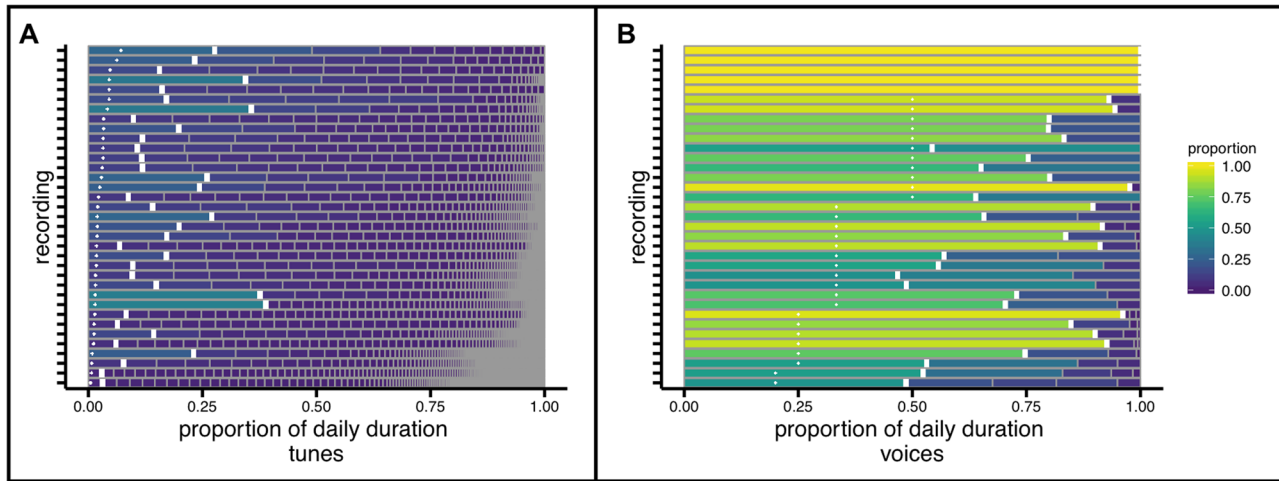


FIGURE 4 Non-uniformly available tune and voice identities in individual daylong recordings of everyday music in infancy. Each row is one recording, segmented into unique identities organized from most-to-least available (left-to-right). The observed proportion of each recording's most available identity (thick white vertical line) exceeded the per-identity proportion expected from uniformly distributed identities (small white +) in daily distributions of (A) tunes and (B) voices. Recordings in each panel are sorted according to the uniform per-identity proportion.

Each recording's observed quantity (total musical seconds in the relevant identity space) and range (number of unique identities) was used to generate its own expected uniform distribution. One-sample discrete Kolmogorov-Smirnov tests (implemented using the *KSGeneral* package in R; Dimitrova et al., 2018) revealed that the observed and expected distributions reliably differed for nearly all individual recordings: tunes (*Median D* = .47, *IQR* = .14, *Range* = .30-.61; 35 recordings $p < .05$) and voices (*Median D* = .32, *IQR* = .28, *Range* = 0-.71; 29 recordings $p < .05$). Thus, the non-uniformity of daily tune and voice distributions appears to be a general property of everyday music in infancy.

Figure 4 shows the striking pattern of non-uniformly available tunes (Figure 4a) and voices (Figure 4b) in individual recordings. Each row is one recording segmented into unique identities. In order to highlight distributional structure, the horizontal extent of a segment represents the proportional availability of its identity in that recording's daily musical seconds. The key observation is that daily recordings are not evenly segmented such that each individual tune or voice is equally available, but rather some identities are very much more available than others.

By definition, non-uniform distributions are comprised of some identities that are more available than others. Fitting curves to the full shape of distributions is not straightforward¹ (Clauset et al., 2009); accordingly, we quantified the “super availability” as the extent to which the most available identity exceeded the per-identity availability expected under a uniform distribution. In Figure 4, this quantity is the difference between the solid white line and the small white + in each recording. One sample Wilcoxon signed rank tests revealed that this “super availability” exceeded zero for tunes (*Median* = .12, *IQR* = .12, *Range* = .02-.37; $V = 630$, $Z = 5.15$, $p < .001$) and for voices

(*Median* = .30, *IQR* = .33, *Range* = 0-.71; $V = 465$, $Z = 4.77$, $p < .001$). Thus, a single voice or tune is very much more available – from .12 to .30 beyond uniform availability – than others in the everyday music of infancy.

3.7 | Corpus scale: Tune and voice identities are not uniformly available

Here, we aggregated data across recordings in order to illustrate the distributional shapes of tune and voice identity durations under various models of the everyday music of infancy. As in the recording scale analysis, we used observed quantities of total daily music and total daily unique identities to generate expected uniform distributions and then compared the shape of this data-driven expectation to the real observed daily distributions of tune and voice identities.

Figure 5 shows the distribution of expected identities (light gray) and observed identities (dark gray) in each identity space. Expected distributions were generated by dividing each recording's total number of musical seconds by its number of unique identities so that each identity inherited the same duration and then these per-identity durations were aggregated across recordings. For example, suppose a hypothetical Recording1 had 40 unique tunes in 2000 musical seconds and a hypothetical Recording2 had 10 unique tunes in 5000 musical seconds. In the expected distribution, Recording1 would contribute 40 tunes of 50 s each and Recording2 would contribute 10 tunes of 500 s each. In the observed distribution, each recording would contribute the real durations associated with each of its unique tunes. This procedure permits a focus on distributional shape, taking into account the reality of the total amounts of music and range of unique identities encountered by infants in their everyday lives. This is the first study to provide empirical estimates of these quantities, and so the resulting

¹“One can, if feeling particularly bold, assert that the distribution follows a power law.... Unfortunately, this method and other variations on the same theme generate significant systematic errors under relatively common conditions...and as a consequence the results they give cannot be trusted” (Clauset et al., 2009, p.5)

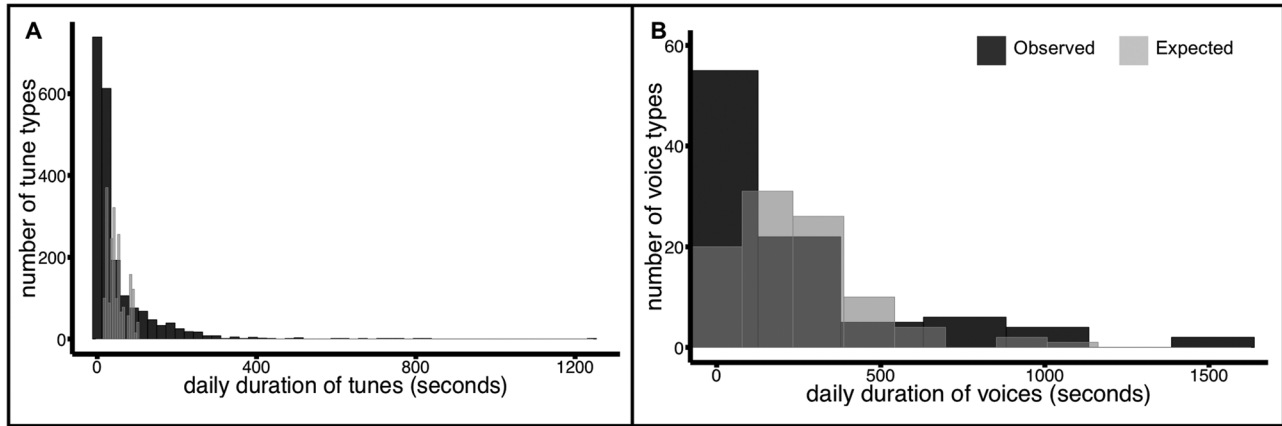


FIGURE 5 Non-uniformly available tune and voice identities in a corpus of everyday music in infancy. The duration of observed identities in this corpus (dark) cumulated to more extreme shorter and longer daily durations than would be expected if uniformly distributed (light) across the daily seconds of (A) tunes and (B) voices. For visualization, histogram bin-widths were determined with respect to each distribution using Scott's rule (Scott, 1979): $h_n = 3.49\sigma n^{-1/3}$.

distributions of raw durations per identity can also constrain future theories of enculturation.

The observed distributions are inconsistent with the expectation that each tune (voice) identity is equally available within a day. Two-sample discrete Kolmogorov-Smirnov tests (R package "Matching"; Sekhon, 2019; 10,000 bootstraps) revealed that the observed and expected distributions reliably differed for tunes ($D(99157) = .34$, $p < .001$) and for voices ($D(21735) = .21$, $p < .001$). The interquartile range of the uniform models suggest that each tune in the everyday music of infancy would be available for 29–63 s per day and each musical voice would be available for 85–296.75 s per day. Instead, per-identity durations outside these ranges were common in the observed everyday music. 63% of individual tunes were available for fewer than 29 daily seconds and 22% for more than 63 daily seconds. Similarly, 46% of particular voices cumulated to fewer than 85 daily seconds and 23% exceeded 296.75 daily seconds.

4 | DISCUSSION

By audio sampling lots of everyday life at home, we discovered that young infants encounter a mix of live, vocal, recorded, and instrumental music. Most instances of this everyday music are recorded and instrumental (e.g., toys) or live and vocal (e.g., caregiver singing). We also discovered that infants encounter multiple tunes and voices per day and that these are distributed unevenly across the day's musical seconds. This first step in describing the musical ecology provides a foundation for future studies of musical enculturation grounded in sensory history.

4.1 | Nine percent of everyday seconds were musical

A broad operationalization of everyday music revealed that nine percent of everyday seconds were musical (151,221 musical seconds

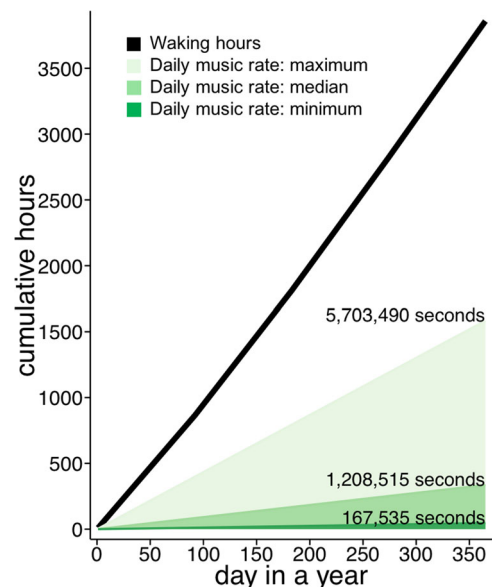


FIGURE 6 Pre-enculturation sensory history: an estimated year of everyday music. Lines show linear extrapolation of the minimum (459 s per day; dark green), median (3,311 s per day; green), and maximum (15,626 s per day; light green) daily musical seconds observed in the present corpus. Waking hours were estimated following Galland et al. (2012).

within the 1,680,351 s of everyday life sampled in this corpus). The observed minimum, median, and maximum musical seconds per day constrain estimates of cumulated musical experience across infants' first post-natal year (Figure 6). Assuming that infants encounter the same amount of music each day (an inference that may be updated upon future multi-day annotations, see below), our observations suggest that between 46 to 1584 h of music are available to infants before soundscape-specific perceptual sensitivities are apparent around their first birthday (Hannon & Trehub, 2005a, 2005b; Lynch & Eilers, 1992; Soley & Hannon, 2010). Properties of this large accumulation of everyday music are what is available to shape musical enculturation.



4.2 | What is frequent in everyday music?

The frequency with which organisms encounter something is a foundational constraint on their opportunities to build knowledge about it (e.g., Hintzman, 1976; Nosofsky, 1988). If one looked to early music learning studies for clues about what musical features are frequently available to young infants, one might expect that infants encounter nearly exclusively recorded instrumental notes and occasionally a recorded voice (Trainor & Corrigan, 2010). From reports about cross-cultural musical universals, one might infer that infants mostly encounter live singing voices (e.g., Mehr et al., 2018; Trehub & Schellenberg, 1995; Trehub & Trainor, 1998; Unyk et al., 1992). Everyday audio sampling licenses conclusions that integrate these two lines of inquiry: we discovered that young infants did encounter a lot of recorded instrumental music and also a lot of live voices. Recorded instrumental music like phrases from toys could provide infants opportunities to build knowledge about instrumental timbre and harmonic structure. Because this music is recorded and thus identical each time it occurs, it may promote discovery of multiple levels of musical structure across exact repetitions (Margulis, 2014). The social quality of live vocal music, in contrast, may encourage heightened attention to its features (Kuhl, 2007; Lebedeva & Kuhl, 2010) as well as back-and-forth practice between caregivers and infants that promotes vocal development (Goldstein et al., 2003) and well-being (Trevvarthen, 1999; Trevvarthen & Malloch, 2000).

Certain individual tunes and voices were also especially available in infants' everyday music. Evidence from adults and children suggests that encountered musical exemplars leave detailed memory traces that are used in subsequent perception (e.g., Corrigan & Trainor, 2010; Creel, 2012; Trainor et al., 2004). Especially available tunes (voices) may be anchors that help infants segment and process their musical soundscape, like familiar names in speech (e.g., Bortfeld et al., 2005). A frequent tune (voice) in everyday music could be easy to discriminate and compare to other tunes (voices) as it occurs across varied contexts, thus helping to structure emerging musical sensitivities (e.g., Creel, 2019a; Smith et al., 2018; Valian & Coulson, 1988). To the extent that the most available exemplars shape perceptual sensitivities more than less available exemplars (e.g., Kelly et al., 2007; Nosofsky, 1988; Oakes & Spalding, 1997), the cumulated availability of some tune and voice identities in everyday music may anchor musical enculturation.

4.3 | How diverse is everyday music?

In researcher-constrained musical tasks, infants typically encounter limited variation (OSF Table 1). In contrast, we discovered that infants encountered diverse musical features in one day. Encountering variable exemplars of a category or categories often accelerates learning by supporting generalization (e.g., Estes & Burke, 1953; Perry et al., 2010). To the extent that each instance of infants' everyday music had typical Western structure (e.g., duple meter), variable instantiations could help infants learn this structure. On average, infants encountered 51 different tunes and 3 different musical voices per day. Such diversity

could help infants extract clusters along dimensions of acoustic variation like those associated with happy (fast tempo, high pitch, major mode) and sad (slow tempo, low pitch, minor mode) emotions (e.g., Schellenberg et al., 2000; see also Rost & McMurray, 2009) and support generalizations of melodic structure (e.g., Loui & Wessel, 2008).

4.4 | How are instances of everyday music distributed?

The combination of instance frequency and diversity yields the distributional shape of sensory histories. The distributional shape of instances in most researcher-constrained studies of infant music learning is uniform (OSF Table 1). In contrast, we discovered that musical tunes and voices were non-uniformly available in everyday music. The distributional shape of instances encountered in many other early ecologies is also non-uniform, including words (Tamis-LeMonda et al., 2017), faces (Jayaraman et al., 2015), and objects (Clerkin et al., 2017). Indeed, non-uniformity appears to be a general property of sensory histories across many scales and domains (e.g., Manaris et al., 2005; Salakhutdinov et al., 2011; Zipf, 1936, 1949). Evidence from many domains suggests that human learners are sensitive to distributional shape (e.g., Clerkin et al., 2017; Griffiths et al., 2007; Oakes & Spalding, 1997; Romberg & Saffran, 2010). The striking universality of non-uniform distributions in everyday ecologies should make their opportunities and challenges central to theories of development.

We speculate that musical enculturation within only one post-natal year may be supported by everyday sensory histories in which musical instances are non-uniformly available. Non-uniform distributions could maximize opportunities to cumulate knowledge from each encountered instance. Each instance is both an encoding and retrieval opportunity. Instances of superavailable identities are likeliest to be encountered in close enough temporal proximity for each to prompt retrieval of the others (e.g., Rovee-Collier, 1995; Rovee-Collier et al., 1980). This strengthened memory would be relatively robust to decay over time and therefore also available for retrieval whenever learners encounter other musical identities. Retrieval can support integration (e.g., Mack et al., 2018; Schlichting & Preston, 2015) and so non-uniform distributions may promote a network of integrated identities. This network will grow more rapidly to the extent that learners attend to and generalize extant knowledge to identities that are infrequently available in their everyday ecology. Non-uniform distributions may provide critical support for such growth precisely because learners encounter these rarer identities in concert with superavailable identities. For example, novel instances often reap greater attentional priority if they are encountered in contexts with familiar instances (e.g., Hunter & Ames, 1988; Hutchinson et al., 2016; Kidd et al., 2014). Novel instances are also more readily integrated into categories learned from distributions with one frequent and many infrequent exemplars (e.g., Carvalho et al., 2021; Navarro, 2013). A growing set of empirical results shows accelerated learning from non-uniform distributions, consistent with these mechanisms for building knowledge (e.g., Casenhiser & Goldberg,



2005; Clerkin et al., 2017; Elio & Anderson, 1984; Hendrickson et al., 2019; Kurumada et al., 2013).

4.5 | How do these insights about everyday music constrain theories of musical enculturation?

Musical enculturation – perceptual sensitivities constrained by soundscape-specific musical structure over time – is widely understood to depend on everyday musical experiences (e.g., Hannon & Trainor, 2007; Trainor & Corrigan, 2010; for additional considerations see Trehub, 2003; Trehub & Hannon, 2006). Prior research suggests multiple patterns of when and how these perceptual sensitivities are observed over the course of development. It is possible that different mechanisms yield these different observed developmental trajectories. We describe three such patterns and then discuss how our research sets the stage to discover which mechanisms of musical enculturation are most likely given the quantitative details of infants' encountered everyday music.

One pattern is that some perceptual sensitivities are observed early. If infants continue to encounter music with soundscape-specific structure (e.g., complex meter, Hannon & Trehub, 2005a, 2005b), then these early-emerging sensitivities persist and may even be strengthened. A mechanism in which everyday music maintains early-emerging sensitivities could drive this pattern ("Maintain"). Another pattern is that some perceptual sensitivities are not observed early but instead emerge later in development. These perceptual sensitivities are more likely to emerge only after infants have accumulated exposure to music with soundscape-specific structure (e.g., scale context, Lynch & Eilers, 1992). A mechanism in which everyday music potentiates subsequent sensitivities could drive this pattern ("Potentiate"). A third pattern is that some perceptual sensitivities emerge after which alternate sensitivities do not (e.g., Hannon & Trehub, 2005b). When infants encounter musical structure inconsistent with their accrued history, then perceptual sensitivities to these alternate structures struggle to emerge. A mechanism in which everyday music entrenches some sensitivities that then impede alternates could drive this pattern ("Entrench"). The extent to which perceptual sensitivities depend on the quantitative details—dose (cumulative frequency) and distribution—of encountered everyday music is as yet unknown for any of the hypothesized "maintain", "potentiate", or "entrench" mechanisms for musical enculturation.

We speculate that pursuing hypotheses about dose (cumulative frequency) and distribution dependencies will advance theories of musical enculturation. For example, perhaps fewer cumulated musical seconds are required to maintain than to potentiate perceptual sensitivities. Perhaps the degree of super-availability in tune distributions covaries with entrenchment strength. Ultimately, we should understand which encountered dose(s) and distribution(s) are most consistent with specific patterns of enculturation. The present discoveries move us a step closer to these long-term goals by constraining likely values for both dose and distribution in early musical ecologies. We now know that infants are likely to encounter between roughly 400 to 16,000 s of music per day. We also now know that there is no sensible average daily

duration per tune but rather that infants encounter many tunes for less than 16 s and a few tunes for more than 100 s per day (Figure 5a; uniform minimum and maximum). Further, we can now consider some sensory histories implausible, like small quantities of music with a single tune or two tunes encountered equally often per day. Our discoveries about everyday music move us away from the doses and distributions that have been prevalent in researcher-constrained tasks (OSF Table 1) and toward theories of experience-dependent change grounded in real experience (Adolph et al., 2018; de Barbaro, 2019; Dahl, 2017; Franck, 2019; Frankenhuys et al., 2019; Rogoff et al., 2018; Smith et al., 2018).

4.6 | Next steps

To build theories of musical enculturation grounded in sensory history, theorists should next augment available histories with everyday snapshots over time and across the world, transcribe everyday music's acoustic space, and manipulate training regimes for both infant and machine learners.

Everyday music over time and across the world. To what extent is everyday music stable across infants' days and cultures? In terms of dose, daily musical quantity is unlikely to be stable given recent reports that early language environments show considerable day-to-day variation (Anderson & Fausey, 2019; d'Apice et al., 2019). Annotations of multiple days per infant will yield insights about the quantity of cumulative musical experiences available to shape enculturation. In terms of distributions, infants could encounter similar music day-to-day. Or, they could encounter a more diverse range of music across days than within a day due to different activities each day. Over the months required to shape soundscape-specific sensitivities, daily distributions could also change slowly with infants' changing motor abilities, interests, and activities (e.g., Fausey et al., 2016). Across the world, it is possible that infants encounter highly similar doses and distributions of daily music, but with acoustics that are specific to their soundscape. Quantifying multiple timescales of sensory histories sampled and aggregated from around the world will advance theories about universals and variation in the developmental timecourse of musical enculturation (Benetti & Costa-Giomi, 2019; Creel, 2019b; Hannon & Trainor, 2007; Mehr et al., 2019; Trainor & Corrigan, 2010). Resources for storing everyday recordings (VanDam et al., 2016) and using standardized manual annotation (Casillas et al., 2017) facilitate collection and aggregation of sensory histories across samples.

The acoustic space of everyday music. The features, tunes, and voices in everyday music structure opportunities for infants to encounter their acoustic properties. Infants are sensitive to many such properties, including melodic contour, tempo, meter, and mode (Trainor & Corrigan, 2010). Researchers should therefore transcribe the individual pitches and their durations of everyday music in order to discover its acoustic space. Many musical properties, such as key, mode, time signature, and infant-directedness, could be derived from transcribed pitches and durations. Among other potential discoveries, acoustic transcription will reveal the extent to which acoustic

properties are shared across the superavailable and the many other identities, governing the likelihood that integrative and contrastive attention and memory mechanisms drive early musical enculturation. Because our corpus is publicly available (Mendoza & Fausey, 2018), other researchers could annotate many properties of the musical sounds in addition to pitch and duration, such as musical genre. Acoustic transcription of our corpus and others will yield insights into many levels of sensory history available to shape musical enculturation.

Manipulating training regimes for infant and machine learners. Testing hypotheses about how everyday doses and distributions of music shape perceptual sensitivities requires manipulating musical experiences at their naturally extended scale. It is not possible to instantiate doses like 46 or more hours of cumulated music, or distributions like a super-available tune paired with very many less available tunes, in experimental protocols lasting minutes-in-a-lab. The value of manipulating experiences at scale is to build theories of developmental change that address the opportunities and challenges encountered by real-infants not lab-infants. One intriguing possibility is that we will discover that there is not one special configuration of encountered music that yields enculturation but rather many pathways through a constrained range of musical experiences that shape development (see Thelen et al., 1996, for related ideas). For example, perhaps many kinds of distributions change perceptual sensitivities but fewer cumulated seconds are required to achieve a particular degree of soundscape-specific behavior if the musical identities are non-uniformly distributed across those seconds. A complementary approach to discovery is to also train models of developmental change with everyday-inspired doses and distributions of music (e.g., Bambach et al., 2016; Ossmy et al., 2018).

4.7 | Conclusion

Models of everyday ecologies focus our collective imagination on the sensory histories available to shape development. Longform audio sampling of infants' everyday lives revealed a plausible sensory history of 55 min of music per day, a quarter of which are live and vocal, with one of the day's 51 tunes and 3 voices superavailable. Such realities about the frequency, range, and distributional shape of encountered music are not discoverable by constraining who is producing what kind of music in what context. Everyday sensory histories are the input to real learners and can now be the input to next-generation theories of musical enculturation.

ACKNOWLEDGMENTS

We thank Catherine Diercks and Christine White for contributing to data collection and we thank 38 undergraduate researchers in the University of Oregon Learning Lab for contributing to annotating recordings. This research was funded in part by a grant from the GRAMMY Museum® to Caitlin M. Fausey.

CONFLICTS OF INTEREST

The authors declare no conflict of interest.

DATA AVAILABILITY STATEMENT

In accordance with family consent, audio recordings and extracted music clips are available on HomeBank (<https://doi.org/10.21415/T5JM4R>; <https://doi.org/10.21415/T47D-5K51>). Study materials, behavioral coding manuals, numerical data, and analysis code are available on Open Science Framework (<https://doi.org/10.17605/osf.io/eb9pw>).

REFERENCES

- Adolph, K. E., Hoch, J. E., & Cole, W. G. (2018). Development (of walking): 15 suggestions. *Trends in Cognitive Sciences*, 22(8), 699–711. <https://doi.org/10.1016/j.tics.2018.05.010>
- Anderson, H., & Fausey, C.M. (2019). Modeling non-uniformities in infants' everyday speech environments [Conference Paper]. In *2019 biennial meeting of the society for research in child development*.
- Aslin, R. N. (2017). Statistical learning: a powerful mechanism that operates by mere exposure. *Wiley Interdisciplinary Reviews: Cognitive Science*, 8(1–2), e1373. <https://doi.org/10.1002/wcs.1373>
- Bambach, S., Crandall, D. J., Smith, L. B., & Yu, C. (2016). Active viewing in toddlers facilitates visual object learning: An egocentric vision approach. In *Proceedings of the 38th Annual Conference of the Cognitive Science Society*. Cognitive Science Society.
- Benetti, L., & Costa-Giomi, E. (2019). Music in the lives of American and Tanzanian infants and toddlers: A daylong sampling [Conference Paper]. In *2019 meeting of the society for music perception and cognition*.
- Bergelson, E., & Aslin, R. N. (2017). Nature and origins of the lexicon in 6-month-olds. *Proceedings of the National Academy of Sciences*, 114(49), 12916–12921. <https://doi.org/10.1073/pnas.1712966114>
- Bergeson, T., & Trehub, S. E. (2002). Absolute pitch and tempo in mothers' songs to infants. *Psychological Science*, 13(1), 72–75. <https://doi.org/10.1111/1467-9280.00413>
- Bortfeld, H., Morgan, J. L., Golinkoff, R. M., & Rathbun, K. (2005). Mommy and me: familiar names help launch babies into speech-stream segmentation. *Psychological Science*, 16(4), 298–304. <https://doi.org/10.1111/j.0956-7976.2005.01531.x>
- Carvalho, P.F., Chen, Ch., & Yu, C. (2021). The distributional properties of exemplars affect category learning and generalization. *Scientific Reports*, 11, 11263. <https://doi.org/10.1038/s41598-021-90743-0>
- Casenhiser, D., & Goldberg, A. E. (2005). Fast mapping between a phrasal form and meaning. *Developmental Science*, 8(6), 500–508. <https://doi.org/10.1111/j.1467-7687.2005.00441.x>
- Casillas, M., Bergelson, E., Warlaumont, A.S., Cristia, A., Soderstrom, M., VanDam, M., Sloetjes, H. (2017). A new workflow for semi-automatized annotations: tests with long-form naturalistic recordings of children's language environments. *Proceedings of Interspeech*, 2017, 2098–2102. <https://doi.org/10.21437/Interspeech.2017-1418>
- Cirelli, L. K., Jurewicz, Z. B., & Trehub, S. E. (2019). Effects of maternal singing style on mother–infant arousal and behavior. *Journal of Cognitive Neuroscience*, 1–8. https://doi.org/10.1162/jocn_a_01402
- Cirelli, L. K., & Trehub, S. E. (2018). Infants help singers of familiar songs. *Music & Science*, 1, 1–11. <https://doi.org/10.1177/2059204318761622>
- Clauset, A., Shalizi, C. R., & Newman, M. E. J. (2009). Power-law distributions in empirical data. *SIAM Review*, 51(4), 661–703. <https://doi.org/10.1137/070710111>
- Clerkin, E. M., Hart, E., Reh, J. M., Yu, C., & Smith, L. B. (2017). Real-world visual statistics and infants' first-learned object names. *Philosophical Transactions of the Royal Society B: biological Sciences*, 372(1711), 20160055. <https://doi.org/10.1098/rstb.2016.0055>
- Conrad, N. J., Walsh, J., Allen, J. M., & Tsang, C. D. (2011). Examining infants' preferences for tempo in lullabies and playsongs. *Canadian Journal of Experimental Psychology/Revue Canadienne de Psychologie Expérimentale*, 65(3), 168–172. <https://doi.org/10.1037/a0023296>



- Corrigall, K. A., & Trainor, L. J. (2010). Musical enculturation in preschool children: acquisition of key and harmonic knowledge. *Music Perception*, 28(2), 195–200. <http://doi.org/10.1525/mp.2010.28.2.195>
- Costa-Giomi, E. (2016). Infant home soundscapes: a case study of 11-month-old twins. *International Perspectives on Research in Music Education*, 70–78.
- Costa-Giomi, E., & Benetti, L. (2017). Through a baby's ears: musical interactions in a family community. *International Journal of Community Music*, 10(3), 289–303. https://doi.org/10.1386/ijcm.10.3.289_1
- Costa-Giomi, E., & Sun, X. (2016). Infants' home soundscape: a day in the life of a family. In J. Bugos (Ed.), *Contemporary Research in Music Learning Across the Lifespan* (pp. 99–108). Routledge.
- Creel, S. C. (2012). Similarity-based restoration of metrical information: different listening experiences result in different perceptual inferences. *Cognitive Psychology*, 65(2), 321–351. <https://doi.org/10.1016/j.cogpsych.2012.04.004>
- Creel, S. C. (2019a). The familiar-melody advantage in auditory perceptual development: parallels between spoken language acquisition and general auditory perception. *Attention, Perception, & Psychophysics*, 81(4), 948–957. <https://doi.org/10.3758/s13414-018-01663-7>
- Creel, S. C. (2019b). Protracted perceptual learning of auditory pattern structure in spoken language. In K.D. Federmeir (Ed.), *The Psychology of Learning and Motivation* (Vol. 71, pp. 67–105). Academic Press.
- Custodero, L. A., & Johnson-Green, E. A. (2003). Passing the cultural torch: musical experience and musical parenting of infants. *Journal of Research in Music Education*, 51(2), 102–114. <https://doi.org/10.2307/3345844>
- Dahl, A. (2017). Ecological commitments: why developmental science needs naturalistic methods. *Child Development Perspectives*, 11(2), 79–84. <https://doi.org/10.1111/cdep.12217>
- d'Apice, K., Latham, R. M., & von Stumm, S. (2019). A naturalistic home observational approach to children's language, cognition, and behavior. *Developmental Psychology*, 55(7), 1414–1427. <https://doi.org/10.1037/dev0000733>
- de Barbaro, K. (2019). Automated sensing of daily activity: a new lens into development. *Developmental Psychobiology*, 61(3), 444–464. <https://doi.org/10.1002/dev.21831>
- Dimitrova, D. S., Kaishev, V. K., & Tan, S. (2018). KSGeneral. R package version 0.1.1. <https://cran.r-project.org/web/packages/KSGeneral/index.html>
- Elio, R., & Anderson, J. R. (1984). The effects of information order and learning mode on schema abstraction. *Memory & Cognition*, 12(1), 20–30. <https://doi.org/10.3758/BF03196994>
- Estes, W. K., & Burke, C. J. (1953). A theory of stimulus variability in learning. *Psychological Review*, 60(4), 276–286. <https://doi.org/10.1037/h0055775>
- Fancourt, D., & Perkins, R. (2018). Maternal engagement with music up to nine months post-birth: findings from a cross-sectional study in England. *Psychology of Music*, 46(2), 238–251. <https://doi.org/10.1177/0305735617705720>
- Fausey, C. M., Jayaraman, S., & Smith, L. B. (2016). From faces to hands: changing visual input in the first two years. *Cognition*, 152, 101–107. <https://doi.org/10.1016/j.cognition.2016.03.005>
- Fausey, C. M., & Mendoza, J. K. (2018). FauseyTrio HomeBank Corpus. <https://doi.org/10.21415/T5JM4R>
- Fenson, L., Dale, P., Reznick, J., Thal, D., Bates, E., Hartung, S., et al., (1993). *The MacArthur Communicative Development Inventories: user's guide and technical manual*. Paul H. Brookes Publishing Co.
- Ford, M., Baer, C. T., Xu, D., Yapanel, U., & Gray, S. (2008). The LENA™ language environment analysis system: audio specifications of the DLP-0121. Retrieved from http://www.lenafoundation.org/wp-content/uploads/2014/10/LTR-03-2_Audio_Specifications.pdf
- Franchak, J. M. (2019). Changing opportunities for learning in everyday life: infant body position over the first year. *Infancy*, 24(2), 187–209. <https://doi.org/10.1111/inf.12272>
- Frankenhuis, W. E., Nettle, D., & Dall, S. R. (2019). A case for environmental statistics of early-life effects. *Philosophical Transactions of the Royal Society B*, 374(1770), 20180110. <https://doi.org/10.1098/rstb.2018.0110>
- Galland, B. C., Taylor, B. J., Elder, D. E., & Herbison, P. (2012). Normal sleep patterns in infants and children: a systematic review of observational studies. *Sleep Medicine Reviews*, 16(3), 213–222. <http://doi.org/10.1016/j.smrv.2011.06.001>
- Goldstein, M. H., King, A. P., & West, M. J. (2003). Social interaction shapes babbling: testing parallels between birdsong and speech. *Proceedings of the National Academy of Sciences*, 100(13), 8030–8035. <https://doi.org/10.1073/pnas.1332441100>
- Goodman, J. C., Dale, P. S., & Li, P. (2008). Does frequency count? Parental input and the acquisition of vocabulary. *Journal of Child Language*, 35(3), 515–531. <https://doi.org/10.1017/S0305000907008641>
- Gottlieb, G. (1991). Experiential canalization of behavioral development: theory. *Developmental Psychology*, 27(1), 4–13. <https://doi.org/10.1037/0012-1649.27.1.4>
- Griffiths, T. L., Steyvers, M., & Tenenbaum, J. B. (2007). Topics in semantic representation. *Psychological Review*, 114(2), 211–244. <https://doi.org/10.1037/0033-295X.114.2.211>
- Hannon, E. E., & Trainor, L. J. (2007). Music acquisition: effects of enculturation and formal training on development. *Trends in Cognitive Sciences*, 11(11), 466–472. <https://doi.org/10.1016/j.tics.2007.08.008>
- Hannon, E. E., & Trehub, S. E. (2005a). Metrical categories in infancy and adulthood. *Psychological Science*, 16(1), 48–55. <https://doi.org/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*, 102(35), 12639–12643. <https://doi.org/10.1073/pnas.0504254102>
- Hendrickson, A. T., Perfors, A., Navarro, D. J., & Ransom, K. (2019). Sample size, number of categories and sampling assumptions: exploring some differences between categorization and generalization. *Cognitive Psychology*, 111, 80–102. <https://doi.org/10.1016/j.cogpsych.2019.03.001>
- Hensch, T. K. (2005). Critical period plasticity in local cortical circuits. *Nature Reviews Neuroscience*, 6(11), 877–888. <https://doi.org/10.1038/nrn1787>
- Hills, T. T., Maouene, J., Riordan, B., & Smith, L. B. (2010). The associative structure of language: contextual diversity in early word learning. *Journal of Memory and Language*, 63(3), 259–273. <https://doi.org/10.1016/j.jml.2010.06.002>
- Hintzman, D. L. (1976). Repetition and memory. In G. H. Bower (Ed.), *The Psychology of Learning and Motivation* (Vol. 10, pp. 47–91). Academic Press.
- Hruschka, D. J., Medin, D. L., Rogoff, B., & Henrich, J. (2018). Pressing questions in the study of psychological and behavioral diversity. *Proceedings of the National Academy of Sciences*, 115(45), 11366–11368. <https://doi.org/10.1073/pnas.1814733115>
- Hunter, M. A., & Ames, E. W. (1988). A multifactor model of infant preferences for novel and familiar stimuli. *Advances in Infancy Research*, 5, 69–95.
- Hutchinson, J. B., Pak, S. S., & Turk-Browne, N. B. (2016). Biased competition during long-term memory formation. *Journal of Cognitive Neuroscience*, 28(1), 187–197. https://doi.org/10.1162/jocn_a_00889
- Ilari, B. (2005). On musical parenting of young children: musical beliefs and behaviors of mothers and infants. *Early Child Development and Care*, 175(7-8), 647–660. <https://doi.org/10.1080/0300443042000302573>
- Ilari, B., & Polka, L. (2006). Music cognition in early infancy: infants' preferences and long-term memory for Ravel. *International Journal of Music Education*, 24(1), 7–20. <https://doi.org/10.1177/0255761406063100>
- Ilari, B., & Sundara, M. (2009). Music listening preferences in early life: infants' responses to accompanied versus unaccompanied singing. *Journal of Research in Music Education*, 56(4), 357–369. <https://doi.org/10.1177/0022429408329107>
- Jayaraman, S., Fausey, C. M., & Smith, L. B. (2015). The faces in infant-perspective scenes change over the first year of life. *PLoS ONE*, 10(5), e0123780. <https://doi.org/10.1371/journal.pone.0123780>

- Kelly, D. J., Quinn, P. C., Slater, A. M., Lee, K., Ge, L., & Pascalis, O. (2007). The other-race effect develops during infancy: evidence of perceptual narrowing. *Psychological Science*, 18(12), 1084–1089. <https://doi.org/10.1111/j.1467-9280.2007.02029.x>
- Kidd, C., Piantadosi, S. T., & Aslin, R. N. (2014). The Goldilocks effect in infant auditory attention. *Child Development*, 85(5), 1795–1804. <https://doi.org/10.1111/cdev.12263>
- Kuhl, P. K. (2007). Is speech learning 'gated' by the social brain? *Developmental Science*, 10(1), 110–120. <https://doi.org/10.1111/j.1467-7687.2007.00572.x>
- Kurumada, C., Meylan, S. C., & Frank, M. C. (2013). Zipfian frequency distributions facilitate word segmentation in context. *Cognition*, 127(3), 439–453. <https://doi.org/10.1016/j.cognition.2013.02.002>
- Lebedeva, G. C., & Kuhl, P. K. (2010). Sing that tune: infants' perception of melody and lyrics and the facilitation of phonetic recognition in songs. *Infant Behavior and Development*, 33(4), 419–430. <https://doi.org/10.1016/j.infbeh.2010.04.006>
- Lee, D. K., Cole, W. G., Golenia, L., & Adolph, K. E. (2018). The cost of simplifying complex developmental phenomena: a new perspective on learning to walk. *Developmental Science*, 21(4), e12615. <https://doi.org/10.1111/desc.12615>
- Loui, P., & Wessel, D. (2008). Learning and liking an artificial musical system: effects of set size and repeated exposure. *Musicae Scientiae*, 12(2), 207–230. <https://doi.org/10.1177/102986490801200202>
- Lynch, M. P., & Eilers, R. E. (1992). A study of perceptual development for musical tuning. *Perception & Psychophysics*, 52(6), 599–608. <https://doi.org/10.3758/BF03211696>
- Mack, M. L., Love, B. C., & Preston, A. R. (2018). Building concepts one episode at a time: the hippocampus and concept formation. *Neuroscience Letters*, 680, 31–38. <https://doi.org/10.1016/j.neulet.2017.07.061>
- Manaris, B., Romero, J., Machado, P., Krehbiel, D., Hirzel, T., Pharr, W., & Davis, R. B. (2005). Zipf's law, music classification, and aesthetics. *Computer Music Journal*, 29(1), 55–69.
- Margulis, E. H. (2014). *On repeat: how music plays the mind*. Oxford University Press.
- Mehr, S. A., Singh, M., Knox, D., Ketter, D. M., Pickens-Jones, D., Atwood, S., Lucas, C., Egner, A., Jacoby, N., Hopkins, E. J., Howard, R. M., Hartshorne, J. K., Jennings, M. V., Simson, J., Bainbridge, C. M., Pinker, S., O'Donnell, T. J., Krasnow, M. M., & Glowacki, L. (2019). Universality and diversity in human song. *Science*, 366(6468), eaax0868. <https://doi.org/10.1126/science.aax0868>
- Mehr, S. A., Singh, M., York, H., Glowacki, L., & Krasnow, M. M. (2018). Form and function in human song. *Current Biology*, 28(3), 356–368. <https://doi.org/10.1016/j.cub.2017.12.042>
- Mehr, S. A., Song, L. A., & Spelke, E. S. (2016). For 5-month-old infants, melodies are social. *Psychological Science*, 27(4), 486–501. <https://doi.org/10.1177/0956797615626691>
- Mendoza, J. K., & Fausey, C. M. (2018). MendozaMusic HomeBank Corpus. <https://doi.org/10.21415/T47D-5K51>
- Mendoza, J. K., & Fausey, C. M. (2019). Everyday Music in Infancy. <https://doi.org/10.17605/osf.io/eb9pw>
- Navarro, D. (2013). Finding hidden types: inductive inference in long-tailed environments. In *Proceedings of the 35th Annual Conference of the Cognitive Science Society*. Cognitive Science Society.
- Nielsen, M., Haun, D., Kärtner, J., & Legare, C. H. (2017). The persistent sampling bias in developmental psychology: a call to action. *Journal of Experimental Child Psychology*, 162, 31–38. <https://doi.org/10.1016/j.jecp.2017.04.017>
- Nosofsky, R. M. (1988). Similarity, frequency, and category representations. *Journal of Experimental Psychology: learning, Memory, and Cognition*, 14(1), 54–65. <https://doi.org/10.1037/0278-7393.14.1.54>
- Oakes, L. M., & Spalding, T. L. (1997). The role of exemplar distribution in infants' differentiation of categories. *Infant Behavior and Development*, 20(4), 457–475. [https://doi.org/10.1016/S0163-6383\(97\)90036-9](https://doi.org/10.1016/S0163-6383(97)90036-9)
- Ossmy, O., Hoch, J.E., MacAlpine, P., Hasan, S., Stone, P., & Adolph, K.E. (2018). Variety wins: soccer-playing robots and infant walking. *Frontiers in Neurorobotics*, 12, 19. <https://doi.org/10.3389/fnbot.2018.00019>
- Pan, B. A., Rowe, M. L., Singer, J. D., & Snow, C. E. (2005). Maternal correlates of growth in toddler vocabulary production in low-income families. *Child Development*, 76(4), 763–782. <https://doi.org/10.1111/1467-8624.00498-i1>
- Perry, L. K., Samuelson, L. K., Malloy, L. M., & Schiffer, R. N. (2010). Learn locally, think globally: exemplar variability supports higher-order generalization and word learning. *Psychological Science*, 21(12), 1894–1902. <https://doi.org/10.1177/0956797610389189>
- Plantinga, J., & Trainor, L. J. (2005). Memory for melody: infants use a relative pitch code. *Cognition*, 98(1), 1–11. <https://doi.org/10.1016/j.cognition.2004.09.008>
- Plantinga, J., & Trainor, L. J. (2009). Melody recognition by two-month-old infants. *The Journal of the Acoustical Society of America*, 125(2), EL58–EL62. <https://doi.org/10.1121/1.3049583>
- Politimou, N., Stewart, L., Müllensiefen, D., & Franco, F. (2018). Music@Home: a novel instrument to assess the home musical environment in the early years. *PLoS One*, 13(4), e0193819. <https://doi.org/10.1371/journal.pone.0193819>
- Rogoff, B., Dahl, A., & Callanan, M. (2018). The importance of understanding children's lived experience. *Developmental Review*, 50, 5–15. <https://doi.org/10.1016/j.dr.2018.05.006>
- Romberg, A. R., & Saffran, J. R. (2010). Statistical learning and language acquisition. *Wiley Interdisciplinary Reviews: cognitive Science*, 1(6), 906–914. <https://doi.org/10.1002/wcs.78>
- Rost, G. C., & McMurray, B. (2009). Speaker variability augments phonological processing in early word learning. *Developmental Science*, 12(2), 339–349. <https://doi.org/10.1111/j.1467-7687.2008.00786.x>
- Rovee-Collier, C. (1995). Time windows in cognitive development. *Developmental Psychology*, 31(2), 147–169. <https://doi.org/10.1037/0012-1649.31.2.147>
- Rovee-Collier, C. K., Sullivan, M. W., Enright, M., Lucas, D., & Fagen, J. W. (1980). Reactivation of infant memory. *Science*, 208(4448), 1159–1161. <https://doi.org/10.1126/science.7375924>
- Roy, B. C., Frank, M. C., DeCamp, P., Miller, M., & Roy, D. (2015). Predicting the birth of a spoken word. *Proceedings of the National Academy of Sciences*, 112(41), 12663–12668. <http://doi.org/10.1073/pnas.1419773112>
- Saffran, J. R., Loman, M. M., & Robertson, R. R. (2000). Infant memory for musical experiences. *Cognition*, 77(1), B15–B23. [https://doi.org/10.1016/S0010-0277\(00\)00095-0](https://doi.org/10.1016/S0010-0277(00)00095-0)
- Salakhutdinov, R., Torralba, A., & Tenenbaum, J. (June). L(2011) earning to share visual appearance for multiclass object detection. In *CVPR 2011* (pp. 1481–1488). IEEE.
- Schellenberg, E. G., Krysciak, A. M., & Campbell, R. J. (2000). Perceiving emotion in melody: interactive effects of pitch and rhythm. *Music Perception: an Interdisciplinary Journal*, 18(2), 155–171. <https://doi.org/10.2307/40285907>
- Schlichting, M. L., & Preston, A. R. (2015). Memory integration: neural mechanisms and implications for behavior. *Current Opinion in Behavioral Sciences*, 1, 1–8. <https://doi.org/10.1016/j.cobeha.2014.07.005>
- Scott, D. W. (1979). On optimal and data-based histograms. *Biometrika*, 66(3), 605–610. <https://doi.org/10.1093/biomet/66.3.605>
- Scott, L. S., Pascalis, O., & Nelson, C. A. (2007). A domain-general theory of the development of perceptual discrimination. *Current Directions in Psychological Science*, 16(4), 197–201. <https://doi.org/10.1111/j.1467-8721.2007.00503.x>
- Sekhon, J. S. (2019). Matching. R package version 4.9-3. <https://cran.r-project.org/web/packages/Matching/index.html>
- Smith, L. B., Jayaraman, S., Clerkin, E., & Yu, C. (2018). The developing infant creates a curriculum for statistical learning. *Trends in Cognitive Sciences*, 22(4), 325–336. <https://doi.org/10.1016/j.tics.2018.02.004>



- Smith, L. B., & Slone, L. K. (2017). A developmental approach to machine learning? *Frontiers in Psychology*, 8, 2124. <https://doi.org/10.3389/fpsyg.2017.02124>
- Soley, G., & Hannon, E. E. (2010). Infants prefer the musical meter of their own culture: a cross-cultural comparison. *Developmental Psychology*, 46(1), 286–292. <https://doi.org/10.1037/a0017555>
- Tamis-LeMonda, C. S., Kuchirko, Y., Luo, R., Escobar, K., & Bornstein, M. H. (2017). Power in methods: language to infants in structured and naturalistic contexts. *Developmental Science*, 20(6), e12456. <https://doi.org/10.1111/desc.12456>
- Thelen, E., Corbetta, D., & Spencer, J. P. (1996). Development of reaching during the first year: role of movement speed. *Journal of Experimental Psychology: human Perception and Performance*, 22(5), 1059–1076. <https://doi.org/10.1037/0096-1523.22.5.1059>
- Trainor, L. J., & Corrigall, K. A. (2010). Music acquisition and effects of musical experience. In M. R. Jones, R. R. Fay, & A. N. Popper (Eds.), *Music Perception* (pp. 89–127). Springer Handbook of Auditory Research.
- Trainor, L. J., Lee, K., & Bosnyak, D. J. (2011). Cortical plasticity in 4-month-old infants: specific effects of experience with musical timbres. *Brain Topography*, 24, 192–203. <https://doi.org/10.1007/s10548-011-0177-y>
- Trainor, L. J., Wu, L., & Tsang, C. D. (2004). Long-term memory for music: infants remember tempo and timbre. *Developmental Science*, 7(3), 289–296. <https://doi.org/10.1111/j.1467-7687.2004.00348.x>
- Trehub, S.E. (2003). Musical predispositions in infancy: an update. In I. Peretz & R.J. Zatorre (Eds.), *The Cognitive Neuroscience of Music* (pp. 3–20). Oxford University Press. <https://doi.org/10.1093/acprof:oso/9780198525202.003.0001>
- Trehub, S.E., & Hannon, E.E. (2006). Infant music perception: domain-general or domain-specific mechanisms? *Cognition*, 100(1), 73–99. <https://doi.org/10.1016/j.cognition2005.11.006>
- Trehub, S. E., & Schellenberg, E. G. (1995). Music: its relevance to infants. *Annals of Child Development*, 11(1), 1–24.
- Trehub, S. E., & Trainor, L. (1998). Singing to infants: lullabies and play songs. In C.K. Rovee-Collier, L.P. Lipsitt, & H. Hayne (Eds.), *Advances in Infancy Research* (Vol., 12, pp. 43–78). Ablex Publishing Corporation.
- Trehub, S. E., Unyk, A. M., Kamenetsky, S. B., Hill, D. S., Trainor, L. J., Henderson, J. L., & Saraza, M. (1997). Mothers' and fathers' singing to infants. *Developmental Psychology*, 33(3), 500–507. <https://doi.org/10.1037/0012-1649.33.3.500>
- Trevarthen, C. (1999). Musicality and the intrinsic motive pulse: evidence from human psychobiology and infant communication. *Musicae Sci. Fall*, 3(Suppl.), 155–215. <https://doi.org/10.1177/10298649000030S109>
- Trevarthen, C., & Malloch, S.N. (2000). The Dance of Wellbeing: defining the Musical Therapeutic Effect. *Nordisk Tidsskrift for Musikterapi*, 9(2), 3–17. <https://doi.org/10.1080/08098130009477996>
- Unyk, A. M., Trehub, S. E., Trainor, L. J., & Schellenberg, E. G. (1992). Lullabies and simplicity: a cross-cultural perspective. *Psychology of Music*, 20(1), 15–28. <https://doi.org/10.1177/0305735692201002>
- U.S. Census Bureau. (2016). American Community Survey, 5-Year Estimates. <https://www.census.gov/data/developers/data-sets/acs-5year.html>
- Valian, V., & Coulson, S. (1988). Anchor points in language learning: the role of marker frequency. *Journal of Memory and Language*, 27(1), 71–86. [https://doi.org/10.1016/0749-596X\(88\)90049-6](https://doi.org/10.1016/0749-596X(88)90049-6)
- VanDam, M., Warlaumont, A. S., Bergelson, E., Cristia, A., Soderstrom, M., De Palma, P., & MacWhinney, B. (May). (2016) HomeBank: an online repository of daylong child-centered audio recordings. In *Seminars in speech and language* (Vol., 37, No. (02), pp. 128–142). Thieme Medical Publishers. <https://doi.org/10.1055/s-0036-1580745>
- Volkova, A., Trehub, S. E., & Schellenberg, E. G. (2006). Infants' memory for musical performances. *Developmental Science*, 9(6), 583–589. <https://doi.org/10.1111/j.1467-7687.2006.00536.x>
- Weiss, M. W., Schellenberg, E. G., Trehub, S. E., & Dawber, E. J. (2015). Enhanced processing of vocal melodies in childhood. *Developmental Psychology*, 51(3), 370–377. <https://doi.org/10.1037/a0038784>
- Werker, J. F., & Hensch, T. K. (2015). Critical periods in speech perception: new directions. *Annual Review of Psychology*, 66, 173–196. <https://doi.org/10.1146/annurev-psych-010814-015104>
- West, M. J., & King, A. P. (1987). Settling nature and nurture into an ontogenetic niche. *Developmental Psychobiology: the Journal of the International Society for Developmental Psychobiology*, 20(5), 549–562. <https://doi.org/10.1002/dev.420200508>
- Wittenburg, P., Brugman, H., Russel, A., Klassmann, A., & Sloetjes, H. (2006). ELAN: a professional framework for multimodality research. In *5th International Conference on Language Resources and Evaluation (LREC 2006)* (pp. 1556–1559).
- Young, S. (2008). Lullaby light shows: everyday musical experience among under-two-year-olds. *International Journal of Music Education*, 26(1), 33–46. <https://doi.org/10.1177/0255761407085648>
- Zipf, G. K. (1936). *The Psychobiology of Language: an introduction to dynamic philology*. Routledge.
- Zipf, G. K. (1949). *Human Behavior and the Principle of Least Effort*. Addison-Wesley Press.

SUPPORTING INFORMATION

Additional supporting information may be found online in the Supporting Information section at the end of the article.

How to cite this article: Mendoza, J. K., & Fausey, C. M. (2021). Everyday music in infancy. *Developmental Science*, 24, e13122. <https://doi.org/10.1111/desc.13122>