Received: 16 May 2022 | Revised: 12 February 2023 | Accepted: 13 February 2023

DOI: 10.1111/desc.13383

Developmental Science WILEY

RESEARCH ARTICLE

Music and language in the crib: Early cross-domain effects of experience on categorical perception of prominence in spoken language

Alan Langus[1] | Natalie Boll-Avetisyan[1] | Sandrien van Ommen[2] | Thierry Nazzi[3]

[1]Cognitive Sciences, Department of Linguistics, University of Potsdam, Potsdam, Germany

[2]Neurosciences Fondamentales, University of Geneva, Geneva, Switzerland

[3]Integrative Neuroscience and Cognition Center, CNRS – Université Paris Cité, Paris, France

Correspondence
Alan Langus and Natalie Boll-Avetisyan, Cognitive Sciences, Department of Linguistics, University of Potsdam, Karl-Liebknecht-Str. 24–25, D-14476 Potsdam, Germany.
Email: langus@uni-potsdam.de and nboll@uni-potsdam.de

Abstract

Rhythm perception helps young infants find structure in both speech and music. However, it remains unknown whether categorical perception of suprasegmental linguistic rhythm signaled by a co-variation of multiple acoustic cues can be modulated by prior between- (music) and within-domain (language) experience. Here we tested 6-month-old German-learning infants' ability to have a categorical perception of lexical stress, a linguistic prominence signaled through the co-variation of pitch, intensity, and duration. By measuring infants' pupil size, we find that infants as a group fail to perceive co-variation of these acoustic cues as categorical. However, at an individual level, infants with above-average exposure to music and language at home succeeded. Our results suggest that early exposure to music and infant-directed language can boost the categorical perception of prominence.

KEYWORDS
categorical perception, cross-domain transfer, infant-directed speech, lexical stress, musical experience, pupillometry

Research Highlights

- 6-month-old German-learning infants' ability to perceive lexical stress prominence categorically depends on exposure to music and language at home.
- Infants with high exposure to music show categorical perception.
- Infants with high exposure to infant-directed language show categorical perception.
- Co-influence of high exposure to music and infant-directed language may be especially beneficial for categorical perception.
- Early exposure to predictable rhythms boosts categorical perception of prominence.

1 INTRODUCTION

Language acquisition research suggests that prosody, that is, the suprasegmental aspects of language that include speech rhythm and prominence, is highly relevant for language acquisition. At birth, human infants are sensitive to prosodic information that allows them to discriminate rhythmically different languages (Gasparini et al., 2021; Mehler et al., 1988) or syllable pairs differing in their stress patterns (Sansavini et al., 1997), pitch contours, or duration/pitch properties (Abboub et al., 2016). Because prosody is aligned with words and syntactic units, it may help infants learn the structural properties of their languages (Morgan & Demuth, 1996; Prieto & Esteve Gibert, 2018). During the first year of life, infants rapidly acquire important prosodic features of their target language, as attested by studies showing cross-linguistic differences in infants' sensitivity to lexical stress (Höhle et al., 2009; Skoruppa et al., 2009) and major prosodic boundaries (Johnson & Seidl, 2008; van Ommen et al., 2020). Moreover, infants start relying on language-specific prosodic cues for segmenting words from speech by about 4–6 months of age (e.g., Jusczyk et al., 1999; Nishibayashi et al., 2015). In the present study, we test the unexplored issue of whether infants learning a language with a lexical stress contrast perceive such a contrast categorically, and investigate whether categorical perception (or: categorization, McMurray, 2022) is modulated by infants' prior experience with music and infant-directed language registers at home.

While prosody is acquired early in development, starting in the womb, and it plays an important role in the initial stages of language development, little is known about how infants mentally represent prosodic regularities. In particular, it is unknown whether infants' prosodic discrimination abilities solely stem from sensitivity to the acoustic distance between prominent and non-prominent vowels, or whether they result from the establishment of prosodic categories, which could be revealed by categorical perception. Categorical perception is a mechanism through which humans and other animals (Green et al., 2020) filter the continuous variability in their environment into discrete categories. Categorical perception has been found across domains, including vision (for faces: Beale & Keil, 1995; facial expressions: Kotsoni et al., 2001; colors: Yang et al., 2016), audition (music: Siegel & Siegel, 1977; environmental sounds: Cutting, 1982) and speech (Liberman et al., 1957). For speech, categorical perception has been observed with adults and infants for phonetic continua of consonants (Eimas et al., 1971; Liberman et al., 1967; Streeter, 1976) and lexical tones (e.g., Hallé et al., 2004; see Tsao & Liu, 2020 for a review). Moreover, categorical perception is attested in the domain of music and non-linguistic sound. Adult listeners have been shown to perceive a continuous variation in sound intensity (Durlach & Braida, 1969), temporal intervals (Burns & Ward, 1974; Sazaki et al., 1998), and tone intervals categorically (Siegel & Siegel, 1977; Locke & Kellar, 1973). Infants have been found to have a categorical perception of musical instruments and timbre in musical stimuli (Trehub & Thorpe, 1989; Trehub et al., 1987, 1990). This suggests that the ability to parse acoustic continua into categorical percepts of prominence in both music and language may be present during the first months of life. Note that in this long history of research on this topic, the categorical perception was originally theorized as a mechanism in which the emergence of categories creates a perceptual warping leading to a widening of the perceptual distance between categories, and a loss of perceptual distance between tokens within the same category. This aspect of the theory has been criticized in recent years (see McMurray, 2022, for a review) most importantly based on evidence that increased experience does not lead to a loss of within-category discriminability but rather to an increased fine-grained sensitivity alongside increased categorization (e.g., Boll-Avetisyan et al., 2018; McMurray et al., 2002; Nixon et al., 2018). While, for these reasons, some authors have proposed to call this phenomenon "categorization," to forego the link to disproved aspects of the original research paradigm (Holt & Lotto, 2010; McMurray, 2022), we will continue to call it "categorical perception" in the present study, in reference to this whole literature and because our current paradigm is less prone to the central criticism directed at classical identification and discrimination studies.

At present, no work has been conducted on the categorical perception of lexical stress, an important issue to explore given that the individual acoustic cues (pitch, intensity, duration) that signal lexical stress also signal prominence at other levels of the linguistic hierarchy (Cutler et al., 1997; Nespor et al., 2008). Moreover, lexical stress varies in magnitude depending on the location of the prominent syllable in the sentence (Nooteboom, 1972). Furthermore, syllables preceding word, phrase, and utterance boundaries are lengthened (Oller, 1973) even when they carry neither lexical stress nor phrasal prominence (Shattuck-Hufnagel & Turk, 1998). Finally, the realization of prominence varies across speech rates, speaking styles, and registers (Mcclean & Tiffany, 1973). Hence, given all this acoustic variability, lexical stress categories might be less straightforwardly defined and hence potentially more difficult to acquire by infants. Accordingly, our study is the first to address the acquisition of lexical stress categories.

Beyond determining whether categorical perception of lexical stress is present early in development, we also investigate whether auditory experience with either music or infant-directed language facilitates such acquisition. Starting with music experience, the possibility that experience with music facilitates infants' prosodic acquisition has to this point been surprisingly neglected in research. The sound structure of spoken language shares many similarities with music, which is best seen when experience in one domain benefits the acquisition (McMullen & Saffran, 2004) or processing (Patel, 2011) in the other domain. For example, in adults, musical experience is correlated with their perception of non-native lexical tones (Wong et al., 2012), consonants varying in voice onset time (Zuk et al., 2013), second language phonology (Slevc & Miyake, 2006), lexical stress (Kolinsky et al., 2009), and rhythmic grouping of speech (Boll-Avetisyan et al., 2016, 2017). Even preschool children trained with various musical activities can show enhanced phoneme and syllable recognition, word and sentence comprehension, as well as linguistic rhythm perception (Degé & Schwarzer, 2011). Early experience with music can even enhance infants' ability to detect temporal violations in speech (Zhao & Kuhl, 2016), phonetic perception (Lebedeva & Kuhl, 2010), word learning (Franco et al., 2022; Papadimitriou et al., 2021), categorical perception of pitch in music intervals (Siegel & Siegel, 1977), and word-level lexical tones (Chen et al., 2020; Wu et al., 2015).

This research area of music and language cross-domain transfer thus demonstrates positive effects of experience with music on the acquisition and processing of a variety of aspects of language. This may suggest that music experience generally enhances auditory perception abilities. An alternative possibility is, however, that at least some of the found associations are mediated by a direct link between music and enhanced prosody perception, which in turn supports the acquisition and processing of words and syntax. Given the shared properties of language and music at the level of rhythm, transfer between the two domains should be particularly expected in the processing of rhythmic aspects of speech including the fine-grained perception of prosodic dimensions of spoken language such as lexical stress. Expanding this question, we investigate in the current study whether early exposure to music can enhance the categorical perception of lexical stress in spoken language.

A second possibility we investigate here is that infants' prosody acquisition benefits from experience with specific infant-directed language registers such as infant-directed speech (IDS) and infant-directed book reading. Infant-directed speech registers are characterized by exaggerated prosody, including higher pitch, greater pitch modulations, and longer pauses than adult-directed speech (e.g., Fernald et al., 1989). Experimental studies indicate that infants find it easier to segment words from infant-directed than adult-directed speech (Thiessen et al., 2005). Also, their ability to learn words is enhanced if novel words are pronounced in an infant-directed register (Graf Estes & Hurley, 2013; Nencheva et al., 2021; Singh et al., 2009). Moreover, higher degrees of exposure to infant-directed speech are associated with stronger language abilities (see Spinelli et al., 2017, for a meta-analysis). Besides experience with spontaneous infant-directed speech, language development seems to also benefit from experience with read-out language directed to them: later language outcomes are predicted by the quantity of joint book reading in infancy (see Dunst et al., 2012, for a meta-analysis). For this reason, we investigate whether infants' individual variability in categorical perception of lexical stress can be explained by prior experience with both music and infant-directed language registers at home.

1.1 | Current study

To test 6-month-old infants' ability to parse acoustic cues co-varying on a continuum into categorical percepts of prominence, we compared trochaic and iambic lexical stress. We chose four instances of the pseudoword "gaba" from a continuum where pitch/intensity contours and segment duration were simultaneously varied from trochaic to iambic lexical stress (see Figure 1b-d). Two of the words had primarily trochaic prominence, and the other two words had primarily iambic prominence, with the acoustic distance between within-category members being identical to the acoustic distance between the two words closest to the category boundary (see Figure 1). During the experiment, infants listened to short trials consisting of four instances of "gaba," with the first three instances providing the context varying across trials (i.e., within-category and between-category context) and the fourth instance providing the test that was identical across trials. We reasoned that a strong prediction for categorical perception would entail infants: (a) discriminating trochaic lexical stress from iambic lexical stress (between-category trials); and (b) not discriminating between acoustically-different syllable pairs where stress is on the same syllable (within-category trials), even when the acoustic differences in the realization of stress between within- and between-category members is comparable. A looser prediction for categorization would entail infants discriminating between-category trials better than within-category trials, because the former encompasses categorization in addition to the standard phonetic discrimination that happens under both conditions (see McMurray, 2022's discussion of Pisoni & Tash, 1974). No-change, standard, trials were also included.

To assess perception, we measured infants' pupil size. Infants' pupils dilate as the result of surprise or increased cognitive load and can therefore provide a non-invasive physiological measure for studying auditory discrimination in a population where overt responses are not possible (Marimon et al., 2022; Wetzel et al., 2016). Finally, to investigate whether categorical perception of lexical stress prominence is modulated by infant-directed language and music, we assessed infants' exposure to language and music through a parental questionnaire detailing their infant's exposure to reading, singing, reciting nursery rhymes, and playing musical instruments, as well as rating how melodically parents talk to their infant and how musical they are. We expected experience with music and infant-related language to benefit categorical perception of lexical stress.

2 | METHODS

2.1 | Participants

We tested 38 monolingual German-learning infants. Of these, seven infants were removed from the analysis due to an insufficient number of good trials. All remaining 31 infants (Mean age: 6 months 20 days; Range: 6 months 0 days to 7 months 2 days; Gender: girls = 17, boys = 14) were born full-term (maximum 3 weeks prior to predicted birth date), had no known visual or auditory problems, and were not at risk for speech and language disorders. Before the experiment, caregivers gave informed consent for participation. Families received monetary compensation and a diploma for participation.

2.2 | Stimuli

We created a 10-step lexical stress continuum of the bisyllabic nonword "gaba" (see Figure 1a) ranging from a trochaic GAba (strong-weak) to an iambic gaBA (weak-strong). The continuum was based on two naturally-produced "gaba" tokens (with GAba positioned on Step 1 and gaBA on Step 8) selected from a set of tokens recorded by a native speaker of German. Lexical stress in the trochaic word was

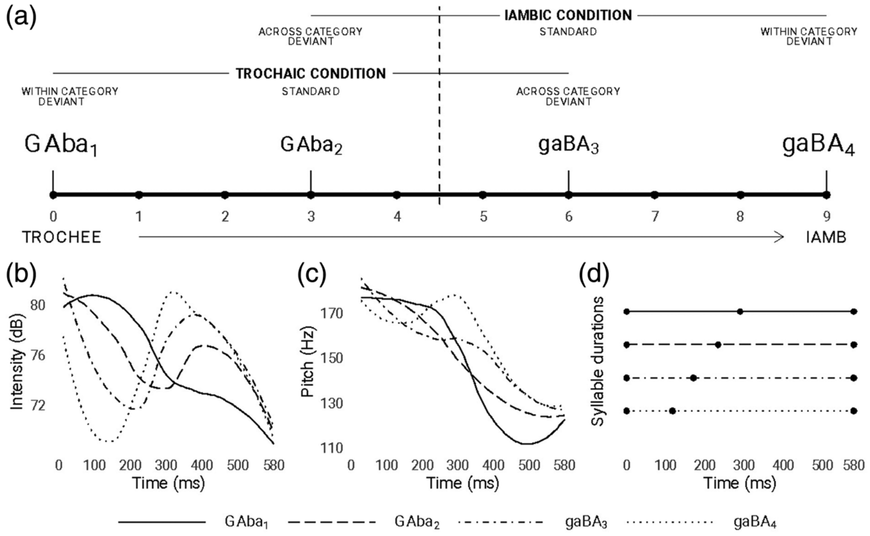

FIGURE 1 Characteristics of the stimuli. (a) Selection of the stimuli from an acoustic continuum, ranging from a Trochee (GAba$_1$) to an Iamb (gaBA$_4$) in equal intervals. (b) Intensity profile of the four words. (C) Pitch profile of the four words. (d) Duration of the syllables of the four words.

TABLE 1 Characteristics of the paradigm. There were two conditions (Trochaic/Iambic) each containing three types of trials: Standard trials where the context and the test items were identical; within-category deviant trials where the context items and the test item constituted a within-category change; and between-category deviants where the change from the context to test constituted a between-category change. The subscripts correspond to the location of the prominence on the continuum depicted in Figure 1.

Condition	Trial Type	Context				Test			
Trochaic	Standard	**GAba$_2$**	ISI	**GAba$_2$**	ISI	**GAba$_2$**	ISI	**GAba$_2$**	Silence
	Within	**GAba$_1$**		**GAba$_1$**		**GAba$_1$**		**GAba$_2$**	
	Between	**gaBA$_3$**		**gaBA$_3$**		**gaBA$_3$**		**GAba$_2$**	
Iambic	Standard	**gaBA$_3$**		**gaBA$_3$**		**gaBA$_3$**		**gaBA$_3$**	
	Within	**gaBA$_4$**		**gaBA$_4$**		**gaBA$_4$**		**gaBA$_3$**	
	Between	**GAba$_2$**		**GAba$_2$**		**GAba$_2$**		**gaBA$_3$**	
1000 ms	baseline	580 ms	400 ms	580 ms	400 ms	580 ms	400 ms	580 ms	1720 ms

signaled primarily through pitch (high-low) and intensity (loud-soft) on the first syllable. In the iambic word, stress was signaled through duration (short-long), intensity (soft-loud) and pitch (low-high) (see Figure 1b-d). Given the natural covariation of the rhythm cues, the continuum was created so that the pitch/intensity contours and segment duration were simultaneously varied in PRAAT in nine acoustically equal intervals (see Supplementary Materials A). For the present experiment, two predominantly trochaic and two iambic items were chosen from the continuum so that the acoustic difference between the items was three steps apart (Figure 1a).

There were two test conditions (trochaic/iambic) and three types of trials (standard, within-category deviant and between-category deviant trials; see Table 1 and Figure 1a). The trials consisted of four “gaba” repetitions: the first three provided the context and the fourth functioned as the test. In the standard trials, the context consisted of the standard iambic or trochaic item. In the within-category deviant trials, the context items were the within-category deviant items. In the between-category deviant trials, the context items were the between-category deviant items. The test item was the same for all trials within a condition (i.e., the standard item). This ensured the change from context to test item was acoustically comparable between within- and between-category deviant trials. The items were separated with 400 ms long pauses and the test item was followed by a 1720 ms long silence, resulting in a total trial length of 5240 ms (see Table 1).

2.3 | Procedure

2.3.1 | Categorical perception experiment

We recorded infants’ pupil size with a Tobii 1750 eye-tracker at 50 Hz. Infants were tested in a room where the lighting was kept constant for all experimental sessions. Infants sat on their caregivers’ lap approximately 65 cm from the eye-tracker display. Caregivers wore blacked-out glasses to avoid the eye-tracker from detecting their gaze and were instructed not to interact with their infant during the experiment. The auditory stimuli were presented over a loudspeaker (65 dB). The experimental session started with a 5-point calibration, after which, infants were tested either in the trochaic or the iambic condition. Each infant received 99 trials in total, one-third per trial type. The start of the trials was contingent on the infant looking at a central fixation continuously for 1000 ms. To draw infants’ gaze to the center

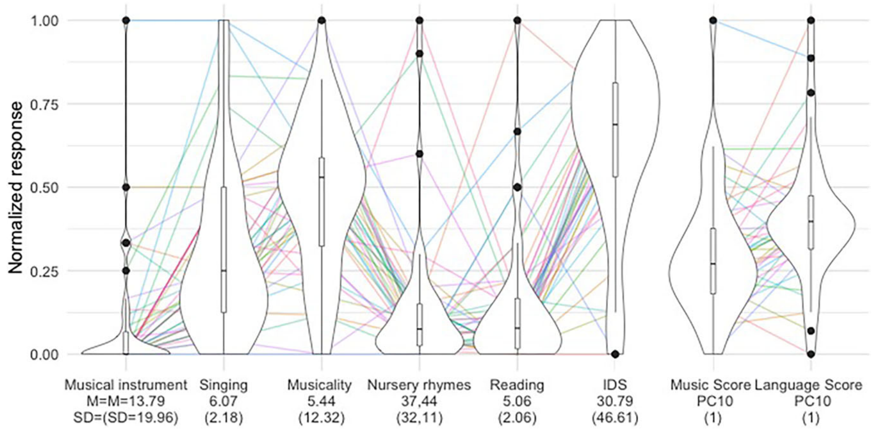

FIGURE 2 Caregivers' responses on the music and infant-directed language questionnaire. The figure depicts the normalized response. The mean and standard deviation values are calculated from untransformed data. The colored lines correspond to individual infants' caregivers.

of the eye-tracker screen, we used a visual attractor identical in shape and luminosity throughout the experiment. The experiment continued until the infant had seen all trials or refused to collaborate further. The experiment was controlled by PsyScope X (http://psy.cns.sissa.it/).

2.3.2 | Musical and language questionnaire

We assessed infants' musical and language experience by asking caregivers to complete a questionnaire detailing their musical and linguistic behavior with their infants. Caregivers were asked how often they: (1) read to their infant; (2) sing to their infant; (3) recite nursery rhymes to their infant; and (4) play musical instruments to their infant. Caregivers were asked to estimate the average time dedicated to each of these activities per day in minutes. Caregivers were also asked to estimate how musical they were and how melodic (i.e., infant-directed speech-like) they spoke to their infant. The latter two questions were asked for each caregiver independently on a 10-point scale (1 corresponding to "not all" and 10 to "very"). Because not all infants who participated in our experiment had two caregivers, not all caregivers indicated their gender or fell into traditional gender roles, we calculated the average response for caregivers for each of the two final questions.

Caregivers' responses are shown in Figure 2. Because the responses to the questionnaire were correlated, we reduced collinearity by performing the principal component analysis (PCA). We applied a predictor reduction via PCA for music- and language-related questions when the answers were correlated. On the one hand, exposure to Musical Instruments, Singing, Nursery Rhymes, and Parents' Musicality were correlated ($r > 0.5$). We call these Musical Questions. On the other hand, exposure to Reading and IDS were correlated ($r > 0.05$) and will here forth be called Language Questions. As music-related questions were not correlated with language-related questions, we conducted two separate PCAs. We used the first principal component (PC) as the predictor, as it accounts for most of the variance of the combined variables. The effect of musical experience (henceforth: Music score) was assessed with a PC, which represented the amount of time dedicated to musical instruments, singing and nursery rhymes as well as how musical caregivers rated themselves (loadings of 0.30, 0.34, 0.44, and 0.64; the four variables being correlated) and captured 53% of the variance contained in these four variables. The effect of infant-directed language experience (henceforth: Language score) was assessed with a PC, which represented the amount of time per day dedicated to reading and how infant-directed speech (IDS)-like caregivers rated their utterances (loadings of 0.85 and 0.73; the two variables being correlated) and captured 72% of the variance contained in these two variables.

2.4 | Pupillary data analysis

To test differences in the pupillary response to the test items, we chose a 2700 ms time window starting 400 ms before test word onset and ending 2300 ms after test word onset, at the end of the trial (Table 1). We only analyzed the pupillary response from the Left eye, as subtle differences between pupil size from the Left/Right eyes might bias the pupillary response. Trials, where more than 25% of eye-tracker samples were missing, were excluded from the analysis, a criterion in accordance with previous pupillometry studies in infants (Marimon et al., 2022). The averaged pupillary response was then linearly interpolated and missing samples at the beginning and at the end of the trial were extrapolated with the median pupil size of the trial. Finally, the data within each trial was linearly detrended. The pupillary response in this time window was baseline-corrected with the average pupil size during the 400 ms silence before test word onset being subtracted from the pupillary response during the whole trial, linearly detrended and smoothed with a cubic spline model. Finally, the pupillary response was averaged over all trials of the same type for cluster based permutation tests. The cleaned data stemmed from 31 infants (16 in the trochaic and 15 in the iambic condition), each infant having on average 33.45 good trials (SD = 17.18), with an average of 10.25 (SD = 5.92) trials for each trial type (i.e., standard, between-category and within-category trials).

To test the time when significant differences between trial types would emerge, we took a cluster-based permutation approach. The conditions were compared at every eye-tracker sample by means of a simple linear regression. Consecutive eye-tracking samples where the linear regression coefficients were significant ($p < 0.05$) were grouped into clusters. Cluster-level statistics were then calculated by summing the t-values within clusters. The reference null

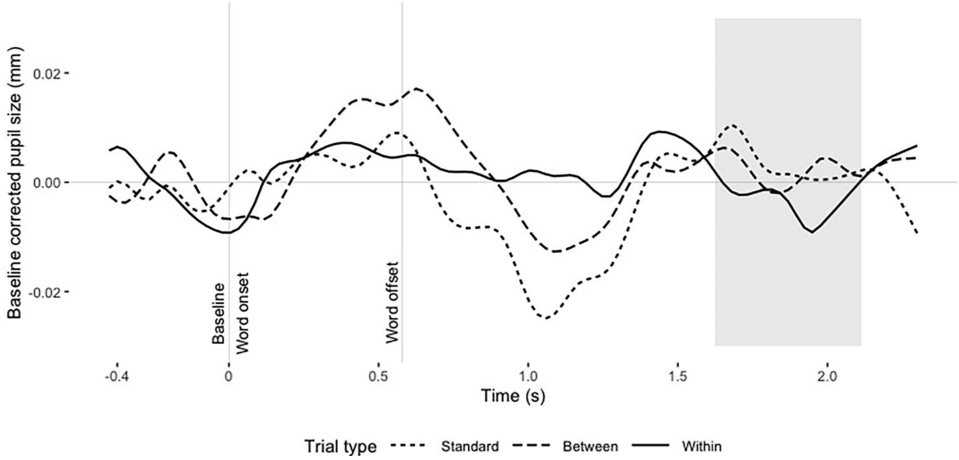

FIGURE 3 Infants' baseline corrected pupillary response (Left pupil) to the test word in the standard, within- and between-category deviant trials showed no significant differences between conditions or trial types. The shaded area corresponds to the time window where we found a three-way significant interaction between Trial Type, Language, and Music Scores.

distribution was calculated using the Monte Carlo method by permuting the mapping between condition labels and experimental data within participants 10k times, each time selecting the cluster with the maximum cluster-level statistics. The number of clusters with higher cluster-level statistics than the actually observed clusters was used to calculate the probability at which the difference between conditions in actual data could be obtained by chance.

3 | RESULTS

3.1 | Categorical perception of lexical stress

Infants' pupillary response to test words is shown in Figure 3. To determine whether and when significant differences between the three conditions emerged, we used cluster-based permutation tests. The model included baseline corrected Pupil Size as the dependent variable. The fixed factors included Trial Type (standard/within-category/between-category) and Condition (trochaic/iambic, cf. Table 1 above). Models including these fixed factors with and without their interactions revealed no eye-tracker samples where the fixed factors or their interactions were significant (all *p*s > 0.05). The failure to observe differences in the pupillary response between standard trials and either within-category or between-category deviant trials suggests that at the group level, discriminating lexical stress patterns is difficult for 6-month-old German-learning infants in our paradigm.

3.2 | Effects of music and infant-directed language scores on categorical perception?

To investigate whether early exposure to various musical activities and infant-directed language influence categorical perception of prominence we used the parental questionnaire scores as correlates for infants' pupillary response at test. The model we used for cluster-based permutation test included baseline corrected Pupil Size (Left eye) as the dependent variable. The fixed factors included the Trial Type (standard/within-category/between-category) as a categorical fixed factor and the Music and Language Scores as continuous and Condition (trochaic/iambic) as categorical between subjects fixed factors. Interactions were included in the model if significant and disregarded otherwise (see Supplementary Materials B). To reduce collinearity all continuous fixed factors were scaled and centered. It had generalized variance-inflation factors (adjusted for degrees of freedom) between 1.056 and 1.988 (Fox & Monette, 1992). We found a three-way interaction between Trial Type, Music and Infant-directed language scores starting at 1620 ms after 4th word onset (duration 480 ms, $T_{SUM} = 42.276$, $p = 0.0102$, see Figure 4).

Posthoc analyses with the mean of the pupil size in the window of interest as dependent variable (1620-2100 ms), the same fixed effect structure as permutation testing and Participant as well as Trial number included as random factors with random intercepts showed a three-way interaction between Trial Type, Music and Language scores ($\beta = 0.022$, $SE = 0.007$, $t = 3.254$, $p = 0.001$) (Supplementary Materials B). Bonferroni-corrected pairwise comparisons of estimated marginal means at the minimum/maximum values of music and Language scores by comparing standard trials to between- and within-category deviant trials revealed: (a) infants' with low music and low infant-directed language experience showed no differences in pupil between trial types; (b) Infants with high music but low infant-directed language experience and infants with high infant-directed language but low musical experience showed larger pupil dilation in standard than between-category trials and no difference between standard and within-category trials; and (c) infants with high music and high infant-directed language experience showed larger pupil dilation in between category than in standard trials and no difference between standard and within-category trials (see Table 2).

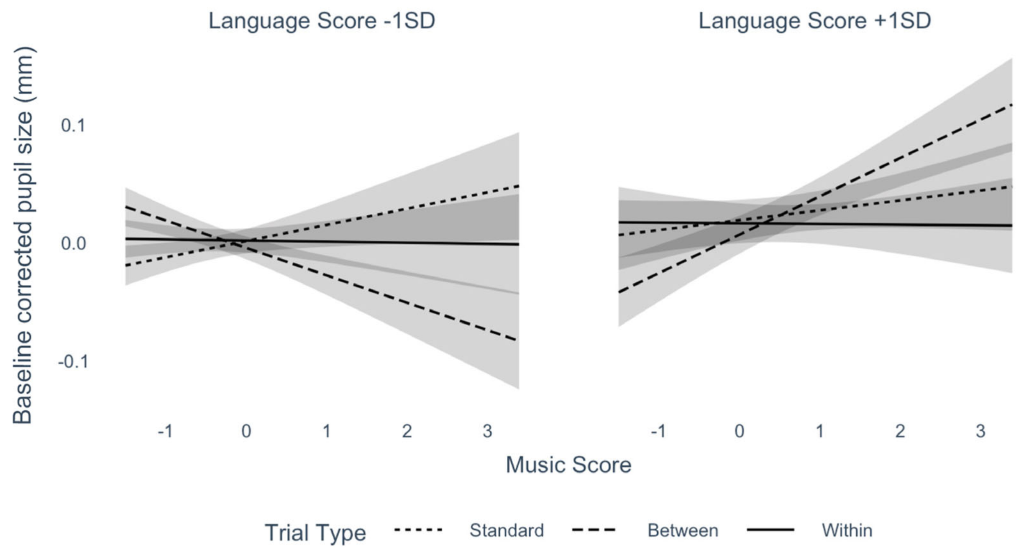

FIGURE 4 Effect of music and language experience at home on infants' pupillary response to the test words in standard, within- and between-category trials. Shaded areas correspond to 95% confidence intervals. The two panels show the effect of music score on pupil size at 1SD from the mean of the language score.

TABLE 2 Pairwise comparisons of estimated marginal means of infants' pupillary response at minimum/maximum values of music and language scores. *p*-values are Bonferroni adjusted for multiple comparisons.

		Standard—Between Category				Standard—Within Category			
Music	IDS	Estimate	SE	T	*p*	Estimate	SE	T	*p*
High	High	−0.160	0.060	−2.648	0.025	0.021	0.061	0.352	1.000
High	Low	0.075	0.028	2.730	0.019	0.027	0.028	0.954	1.000
Low	High	0.131	0.058	2.267	0.024	−0.007	0.058	−0.122	0.903
Low	Low	−0.011	0.009	−1.215	0.225	−0.007	0.009	−0.713	0.476

4 | DISCUSSION

The present study investigated how German-learning infants perceive speech prosody under conditions of acoustic variability, testing whether they perceive lexical stress, a fundamental property of their native language prosodic system, categorically. It further investigated whether early exposure to music and infant-directed language at home influences such categorical perception. We reasoned that if infants perceive prominence categorically, they should (a) discriminate between words carrying word-initial (trochaic) lexical stress versus word-final (iambic) lexical stress; and (b) fail to discriminate between two different word items that both carry either initial or final lexical stress, even when acoustic distance is matched between the two conditions. Under a looser definition of categorization during perception, between-category discrimination should be more pronounced than within-category discrimination. We also hypothesized that experience with a variety of musical activities and exposure to infant-directed language registers at home should boost infants' ability to perceive prominence as categorical.

Even though categorical perception of phonemic continua has been shown at birth and infants can perceive non-linguistic analogs of speech as categorical already by 2 months of age (Jusczyk et al., 1977), the present experiment did not find evidence for consistent categorical perception of prominence in 6-month-old infants as a group. That is, infants' pupillary response at test showed no group-level differences between standard, between- and within-category trials. Given how prevalent categorical perception is in human adults and infants, it is important to consider possible reasons for infants' group-level failure in our experiment to show evidence for categorical perception of prominence.

On the one hand, it is conceivable that the infants' group-level failure in the present experiment is caused by the experimental paradigm being too difficult. In our paradigm, infants were provided with only four instances of the word "gaba," which may be insufficient for them to build strong enough representations of the prominence patterns in the context phase of the trial. Furthermore, the failure to observe systematic pupillary dilation in standard, between- and within-category trials may be caused by the highly repetitive nature of the paradigm where only three types of trials are repeated more than 30 times each. Note however that such a paradigm of 4-item trials has mainly been successfully used in ERP studies to establish discrimination and categorization abilities (e.g., Dehaene-Lambertz & Baillet, 1998), and less frequently so for measuring pupillary responses (Hochmann & Papeo, 2014). Hence, it is possible that it is the combination of using this presentation paradigm and measuring pupillary response that reduces sensitivity of our procedure, leading to the group null result. As a consequence, it is possible that the infants in our paradigm failed to show robust categorical perception because of pragmatic factors associated with the design of the experiment.

Alternatively, it is also possible that the lack of group-level differences in infants' pupillary responses to the three types of trials represent a genuine difficulty to perceive lexical stress in spoken language categorically. While many studies have investigated young infants' ability to perceive prominence in spoken language (for an overview see Bhatara et al., 2018; Langus et al., 2017), many of these studies have relied on stimuli that provided straightforward cues to prominence such as high versus low pitch. However, in realistic linguistic input, the acoustic cues signaling prominence—that is, pitch, intensity, and duration—co-vary on a continuum. For example, in many languages, including German, lexical stress and phrasal prominence is signaled by multiple acoustic cues at once. Furthermore, the realization of prominence does not only depend on the rhythmic structure of spoken language, but also on the location of prominence in sentences as well as on speech rate and register. Our results may therefore suggest that perceiving the co-variation of acoustic cues in spoken language as categorical is difficult for young infants. Note, however, that to control acoustic distance between the between- and within category conditions, acoustic distance for between-category words was smaller than it was in Höhle et al. (2009), in which the trochaic/iambic contrast was based on a natural acoustic distance that was twice as large compared to the acoustic distances between stimuli in the present study. Thus, group failure here could be due to both reduced acoustic distance or increased variability. As a consequence, there might be interindividual variability in categorical perception of lexical stress, which might depend on developmental level, individual abilities, and as explored here, properties of the input received prior to testing, in the home environment.

4.1 | Effects of experience with music and infant-directed language on categorical perception of prominence

We show that early exposure to various musical and linguistic activities can benefit infants in perceiving acoustic prominence categorically. The amount and quality of language input play an important role in language acquisition (Hoareau et al., 2019; Weisleder & Fernald, 2013). Also, early exposure to music can boost infants' language abilities (Franco et al., 2022; Lebedeva & Kuhl, 2010; Papadimitriou et al., 2021; Zhao & Kuhl, 2016). Because infants experience language and music predominantly at home, we asked caregivers to estimate how much time they dedicate to playing instrumental music, singing, reciting nursery rhymes, and reading to their infant. We also asked how musical they rate themselves to be and how melodically they talk to their infant. The correlations between caregivers' responses suggest that the answers were patterned into two categories. Caregivers who rated themselves as highly musical were also more likely to play instrumental music, recite nursery rhymes, and sing to their infant. Conversely, caregivers who spoke in infant-directed speech also read more to their infants. To account for these correlations in our analysis we used the first principal component of the music and language scores in the analysis.

While the 6-month-olds as a group failed to perceive prominence as categorical in our study, we did find evidence for categorical perception when considering infants' experience with music and infant-directed language registers at home. Infants who had on average low music and infant-directed language scores showed no evidence for categorical perception. In contrast, infants with either only high music or only high infant-directed language scores showed significantly larger pupil dilations in standard trials than in between-category trials, but no difference in pupil size between standard and within-category trials. Finally, the combination of high music and high infant-directed language scores showed the expected effect, with infants' pupils dilating more when they received between-category trials, indicating their surprise about the change in category, than when they received standard and within-category trials. Note that we did not find evidence for or against discrimination in the within-category trials. This might support our strong categorical perception prediction. By this, our results show that a combination of various musical activities and infant-directed speech and reading exposure can facilitate the perception of prosody in spoken language by enhancing infants' ability to map perceived pitch, intensity, and duration cues co-varying on a continuum onto lexical stress categories of trochee versus iamb. Infants with low music and language-related experience clearly fail to categorically perceive the continuum (showing in fact no ability to discriminate even the trochaic vs. iambic items), while infants with high music and high language-related experience clearly perceive the continuum categorically.

Importantly, infants' pupillary response at test suggests that the co-influence of high music and infant-directed language exposure at home may lead to differences in categorical perception when compared to infants who only had high exposure in either one of the two domains. Only infants with high exposure in both domains showed the expected effects with larger pupil dilations in response to deviants from a different category when compared to within-category or standard trials. Infants who had high exposure in only one of the two domains showed larger pupil dilation in response to the standard trials than in response to between category trials, which requires explanation. Previous studies have shown that a pupillary response can reflect a number of different processes, including pupil dilation associated with surprise caused by the novelty (as found in most studies, which justified our prediction of larger dilation for between-category over standard trials), but also the recognition and processing of familiar stimuli

(Kafkas & Montaldi, 2012). Note that such familiarity/novelty effects are found in developmental science using other methods, such as the behavioral head turn preference procedure (e.g., Bosch et al., 2013; Goyet et al., 2013) or ERPs (e.g., Kooijman et al., 2013; Männel & Friederici, 2013). Some researchers have proposed that novelty effects in behavioral tasks represent more advanced processing levels than familiarity effects (Hunter & Ames, 1988). The present results are therefore consistent with the idea that those infants who had high exposure to both music and infant-directed language at home had established strong categorical boundaries between trochaic and iambic lexical stress, thus showing a robust pupil dilation as the result of surprise when hearing between-category changes. In contrast, infants who had high experience with only music or only infant-related language, appear to have established weaker category boundaries, as attested by them showing a familiarity effect of engaging more when hearing again the standard word belonging to the same category.

Previous studies have shown that adult listeners can perceive acoustic cues varying on a continuum as categorical with musical and non-linguistic sounds. For example, adult listeners can perceive sound amplitude (Durlach & Braida, 1969), pitch intervals (Sazaki et al., 1998) as well as durational intervals between individual notes categorically (Locke & Kellar, 1973; Siegel and Siegel, 1977). Even infants can perceive temporal, tonal, and timber intervals in musical stimuli categorically (Trehub & Thorpe, 1989; Trehub et al., 1987, 1990). The present results add to this picture by showing that categorical perception of linguistic prominence signaled by the co-variation of pitch, intensity, and duration may be more difficult to acquire, and benefit from experience with acoustically more predictable rhythms in music and infant-directed language registers. The correlation between how much caregivers read and how melodically they speak to their infant, therefore, suggests that the exaggerated pronunciation and reduced acoustic variation observed in infant-directed speech may be necessary for young infants to classify acoustic cues into categorical percepts of prominence (Ludusan et al., 2016).

Finally, our results suggest that the perception of speech prosody is facilitated by cross-domain effects of musical experience. We found that more experience with singing, musical instrument sounds, nursery rhymes, and parents' musicality were associated with infants' ability to perceive linguistic prominence as categorical. Even though songs and nursery rhymes contain spoken language, they are characterized by music-like regular rhythms and patterns in caregivers' responses as separate from exposure to reading and IDS. This may suggest that at least some of the benefits of musical experience observed in the present experiment are explained by the predictable nature of rhythm in music-related activities at home. Alternatively, our results may also suggest that categorical perception of linguistic prominence may be linked to domain-general perceptual processing of sound, with music-related signals boosting infants' ability to perceive prosodic cues as categorical through exposure to acoustic continua in diverse auditory domains. Furthermore, because the present experiment only looked at categorical perception of linguistic prominence, it is not possible at present to say whether experience with infant-related language could also boost categorical perception of prominence in music. However, the results do suggest that the ability to parse acoustic continua into categorical percepts may rely on a domain-general network for auditory processes. Variable experiences with more and less predictable rhythm structures in the domains of music and language seem to facilitate the establishment of reliable prosodic categories It will be interesting for future studies to explore if cross-domain transfer also benefits the perception of music: that is, whether language-related experience factors enhance the categorical perception of musical categories. Few studies (e.g., Bhatara et al., 2015) have, so far, looked at this direction of the link between music and language, but theories of transfer between the two domains would benefit from evidence that the transfer is bidirectional (see Asaridou & McQueen, 2013 for a discussion).

5 | CONCLUSION

Infants are believed to find structure in complex auditory signals by relying on categorical perception, a domain-general and potentially innate perceptual mechanism. However, the prosody in spoken language and rhythm in music unfold under conditions of considerable acoustic variability. Our results show that 6-month-old infants can have difficulties perceiving prominence signaled by the co-variation of pitch, intensity and duration as categorical in short rhythmic sequences. However, early exposure to music and infant-directed language registers at home seems to boost the categorical perception of prominence in speech. While the parsing of acoustic continua into categorical percepts of prominence may therefore be domain-general, it develops during the first year of life and depends on early exposure to music and infant-directed speech. Assessment of experience with language and music in the crib is therefore not only informative about the domain of perceptual processing but also reveals how experience shapes the perception of speech prosody during the first year of life.

ACKNOWLEDGMENTS

This work was supported by: Two Agence Nationale de la Recherche—Deutsche Forschungsgemeinschaft grants awarded to Barbara Höhle/Thierry Nazzi [grant numbers #09-FASHS-018, HO-1960/14—1] and to Ranka Bijeljac-Babic/Barbara Höhle [grant numbers HO-1960/15—1, ANR-13-FRAL-0010]. By the European Union's Horizon 2020 Individual Marie-Curie Fellowships to Alan Langus [grant number 748909; "RHYTHMSYNC"] and Sandrien van Ommen [grant number 892890; "RhythmicPrediction"]. By the Deutsche Forschungsgemeinschaft (DFG, German Research Foundation) funding for the Sonderforschungsbereich (Collaborative Research Centre) "Limits of Variability in Language" (grant number 317633480, SFB 1287, collaboration between projects C03 and C07).

Open access funding enabled and organized by Projekt DEAL.

CONFLICT OF INTEREST STATEMENT

The authors declare no conflict of interest.

DATA AVAILABILITY STATEMENT

The data and the stimuli are available from corresponding authors upon reasonable request.

ETHICS APPROVAL STATEMENT

The research reported in this manuscript was approved by the Ethics Committee of the University of Potsdam (NR50/2019).

ORCID

Alan Langus https://orcid.org/0000-0003-3424-8162
Natalie Boll-Avetisyan https://orcid.org/0000-0001-5446-946X
Sandrien van Ommen https://orcid.org/0000-0002-8311-2774
Thierry Nazzi https://orcid.org/0000-0002-4378-3661

REFERENCES

Abboub, N., Boll-Avetisyan, N., Bhatara, A., Höhle, B., & Nazzi, T. (2016b). An exploration of rhythmic grouping of speech sequences by French- and German- learning infants. *Frontiers in Human Neuroscience*, *10*, 292.

Asaridou, S. S., & McQueen, J. M. (2013). Speech and music shape the listening brain: Evidence for shared domain-general mechanisms. *Frontiers in Psychology*, *4*, 321.

Beale, J. M., & Keil, F. C. (1995). Categorical effects in the perception of faces. *Cognition*, *10*(3), 217–239.

Bhatara, A., Boll-Avetisyan, N., Höhle, B., & Nazzi, T. (2018). Early sensitivity and acquisition of prosodic patterns at the lexical level. In P., Prieto, & N., Esteve-Gibert (Eds.), *The development of prosody in first language acquisition*. John Benjamins Trends in Language Acquisition Research, 23.

Bhatara, A., Yeung, H. H., & Nazzi, T. (2015). Foreign language learning in French speakers is associated with rhythm perception, but not with melody perception. *Journal of Experimental Psychology: Human Perception and Performance*, *41*(2), 277–282.

Boll-Avetisyan, N., Bhatara, A., & Höhle, B. (2017). Effects of musicality on the perception of rhythmic structure in speech. *Laboratory Phonology*, *8*(1), 9.

Boll-Avetisyan, N., Bhatara, A., Unger, A., Nazzi, T., & Höhle, B. (2016). Effects of experience with L2 and music on rhythmic grouping by French listeners. *Bilingualism: Language and Cognition*, *19*(5), 971–986.

Boll-Avetisyan, N., Nixon, J. S., Lentz, T. O., Liu, L., van Ommen, S., Cöltekin, C., & van Rij, J. (2018). Neural response development during distributional learning. In *Proceedings of Interspeech 2018* (pp. 1432–1436). ISCA. https://doi.org/10.21437/Interspeech.2018-2072

Bosch, L., Figueras, M., Teixido, M., & Ramon-Casas, M. (2013). Rapid gains in segmenting fluent speech when words match the rhythmic unit: Evidence from infants acquiring syllable-timed languages. *Frontiers in Psychology*, *4*, 106.

Burns, E., & Ward, W. D. (1974). Categorical Perception of Musical Intervals. *The Journal of the Acoustical Society of America*, *55*, 456–456.

Chen, S., Zhu, Y., Wayland, R., & Yang, Y. (2020). How musical experience affects tone perception efficiency by musicians of tonal and non-tonal speakers? *PloS one*, *15*(5), e0232514. https://doi.org/10.1371/journal.pone.0232514

Cutting, J. E. (1982). Plucks and bows are categorically perceived, sometimes. *Perception & Psychophysics*, *31*, 462–476.

Cutler, A., Dahan, D., & Donselaar, W. (1997). Prosody in the comprehension of spoken language: A literature review. *Language and Speech*, *40*(2), 141–201.

Degé, F., & Schwarzer, G. (2011). The effect of a music program on phonological awareness in preschoolers. *Frontiers in Psychology*, *2*, 124.

Dehaene-Lambertz, G., & Baillet, S. (1998). A phonological representation in the infant brain. *NeuroReport*, *9*, 1885–1888.

Dunst, C. J., Simkus, A., & Hamby, D. W. (2012). Relationship between age of onset and frequency of reading and infants' and toddlers' early language and literacy development. *Center for Early Literacy Learning*, *5*(3), 1–10.

Durlach, N. I., & Braida, L. D. (1969). Intensity perception: I. Preliminary theory of intensity resolution. *Journal of the Acoustical Society of America*, *46*(2, Pt. 2), 372–383.

Eimas, P. D., Siqueland, E. R., Jusczyk, P., & Vigorito, J. (1971). Speech perception in infants. *Science*, *171*(3968), 303–306.

Fox, J., & Monette, G. (1992). Generalized collinearity diagnostics. *JASA*, *87*, 178–183.

Franco, F., Suttora, C., Spinelli, M., Kozar, I., & Fasolo, M. (2022). Singing to infants matters. *Journal of Child Language*, *49*(3), 552–577.

Fernald, A., Taeschner, T., Dunn, J., Papousek, M., de Boysson-Bardies, B., & Fukui, I. (1989). A cross-language study of prosodic modifications in mothers' and fathers' speech to preverbal infants. *Journal of child language*, *16*(3), 477–501.

Gasparini, L., Langus, A., Tsuji, S., & Boll-Avetisyan, N. (2021). Quantifying the role of rhythm in infants' language discrimination abilities. *Cognition*, *213*, 104757.

Goyet, L., Nishibayashi, L-L., & Nazzi, T. (2013). Early syllabic segmentation of fluent speech by infants acquiring French. *PlosOne*, *8*(11), e79646.

Graf Estes, K., & Hurley, K. (2013). Infant-Directed Prosody Helps Infants Map Sounds to Meanings. *Infancy*, *18*(5), 797–824.

Green, P. A., Brandley, N. C., & Nowicki, S. (2020). Categorical perception in animal communication and decision-making, *Behavioral Ecology*, *31*(4), 859–867.

Hallé, P. A., Chang, Y. C., & Best, C. T. (2004). Identification and discrimination of Mandarin Chinese tones by Mandarin Chinese vs. French listeners. *Journal of Phonetics*, *32*(3), 395–421.

Hoareau, M., Yeung, H. H., & Nazzi, T. (2019). Infants' statistical word segmentation in an artificial language is linked to both parental speech input and reported production abilities. *Developmental Science*, *22*(4), e12803.

Hochmann, J.-R., & Papeo, L. (2014). The Invariance Problem in Infancy: A Pupillometry Study. *Psychological Science*, *25*(11), 2038–2046.

Holt, L. L., & Lotto, A. J. (2010). Speech perception as categorization. *Attention, Perception, & Psychophysics*, *72*(5), 1218–1227.

Höhle, B., Bijeljac-Babic, R., Herold, B., Weissenborn, J., & Nazzi, T. (2009). Language-specific prosodic preferences during the first year of life: Evidence from German and French infants. *Infant Behavior and Development*, *32*(3), 262–274.

Hunter, M. A., & Ames, E. W. (1988). A multifactor model of infant preferences for novel and familiar stimuli. In L. P., Lipsitt (Eds.), *Advances in infancy research* (Vol., 5). Ablex.

Johnson, E. K., & Seidl, A. (2008). Clause segmentation by 6-month-olds: A crosslinguistic perspective. *Infancy*, *13*(5), 440–455.

Jusczyk, P. W., Houston, D. M., & Newsome, M. (1999). The beginnings of word segmentation in english-learning infants. *Cognitive Psychology*, *39*(3-4), 159–207.

Jusczyk, P. W., Rosner, B. S., & Cutting, J. E. et al. Categorical perception of nonspeech sounds by 2-month-old infants. *Perception & Psychophysics*, *21*, 50–54.

Kafkas, A., & Montaldi, D. (2012). Familiarity and recollection produce distinct eye movement, pupil and medial temporal lobe responses when memory strength is matched. *Neuropsychologia*, *50*(13), 3080–3093.

Kooijman, V., Junge, C., Johnson, E. K., Hagoort, P., & Cutler, A. (2013). Predictive brain signals of linguistic development. *Frontiers in Psychology*, *4*, 25.

Kotsoni, E., de Haan, M., & Johnson, M. H. (2001). Categorical perception of facial expressions by 7-month-old infants. *Perception*, *30*(9), 1115–1125.

Kolinsky, R., Cuvelier, H., Goftry, V., Peretz, I., & Morais, J. (2009). Music training facilitates lexical stress processing. *Music Perception*, *22*, 235–246.

Langus, A., Mehler, J., & Nespor, M. (2017). Rhythm in language acquisition. *Neuroscience and Biobehavioral Reviews*, *81*(Pt B), 158–166.

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 & Development*, *33*(4), 419–430.

Liberman, A. M., Cooper, F. S., Shankweiler, D. P., & Studdert-Kennedy, M. (1967). Perception of the speech code. *Psychological Review*, *74*(6), 431–461.

Liberman, A. M., Harris, K. S., Hoffman, H. S., & Griffith, B. C. (1957). The discrimination of speech sounds within and across phoneme boundaries *Journal of Experimental Psychology*, *54*(5), 358–368.

Locke, S., & Kellar, L. (1973). Categorical perception in a non-linguistic mode. *Cortex: A Journal Devoted to the Study of the Nervous System and Behavior*, *9*(4), 355–369.

Ludusan, B., Cristia, A., Martin, A., Mazuka, R., & Dupoux, E. (2016). Learnability of prosodic boundaries. *JASA*, *140*(2), 1239.

Marimon, M., Höhle, B., & Langus, A. (2022). Pupillary entrainment reveals Individual variability in cue-weighting in 9-month-old German-learning infants. *Cognition*, *224*, 105054.

Mcclean, M. D., & Tiffany, W. R. (1973). The acoustic parameters of stress in relation to syllable position, speech loudness and rate. *Language and Speech*, *16*(3), 283–290.

McMullen, E., & Saffran, J. R. (2004). Music and language. *Music Perception*, *21*(3), 289–311.

McMurray, B. (2022). The myth of categorical perception. *The Journal of the Acoustical Society of America*, *152*(6), 3819–3842.

McMurray, B., Tanenhaus, M., & Aslin, R. N. (2002). Gradient effects of within-category phonetic variation on lexical access. *Cognition*, *86*(2), B33–B42. https://doi.org/10.1016/S0010-0277(02)00157-9

Männel, C., & Friederici, A. D. (2013). Accentuate or repeat? Brain signatures of developmental periods in infant word recognition. *Corte*, *49*, 2788–2798.

Mehler, J., Jusczyk, P., Lambertz, G., Halsted, N., Bertoncini, J., & Amiel-Tison, C. (1988). A precursor of language acquisition in young infants. *Cognition*, *29*(2), 143–178.

Morgan, J. L., & Demuth, K. (1996). *Signal to syntax: Bootstrapping from speech to grammar in early acquisition*. Lawrence Erlbaum Associates, Inc.

Nencheva, M. L., Piazza, E. A., & Lew-Williams, C. (2021). The moment-to-moment pitch dynamics of child-directed speech shape toddlers' attention and learning. *Developmental Science*, *24*(1), e12997.

Nespor, M., van de Vijver, R., Schraudolf, H., & Donati, C. (2008). Different phrasal prominence realization in VO and OV languages. *Lingue e Linguaggio*, *7*, 1–28.

Nishibayashi, L.-L., Goyet, L., & Nazzi, T. (2015). Early Speech Segmentation in French-learning Infants: Monosyllabic Words versus Embedded Syllables. *Language and Speech*, *58*(3), 334–350.

Nixon, J. S., Boll-Avetisyan, N., Lentz, T. O., van Ommen, S., Keij, B., Çöltekin, Ç., Liu, L., & van Rij, J. (2018). Short-term exposure enhances perception of both between- and within-category acoustic information. *Proceedings of the 9th International Conference on Speech Prosody Speech prosody* (Vol. 2018, pp. 114–118). https://doi.org/10.21437/SpeechProsody.2018-23

Nooteboom, S. G. (1972). *Production and perception of vowel duration*. Doctoral dissertation, Utrecht University.

Oller, D. K. (1973). The effect of position in utterance on speech segment duration in English. *JASA*, *54*, 1235–1247.

Papadimitriou, A., Smyth, C., Politimou, N., Franco, F., & Stewart, L. (2021). The impact of the home musical environment on infants' language development. *Infant Behavior and Development*, *65*, 101651.

Patel, A. D. (2011). Why would musical training benefit the neural encoding of speech? *Frontiers in Psychology*, *2*, 1–14.

Pisoni, D. B., & Tash, J. (1974). Reaction times to comparisons within and across phonetic categories. *Perception & Psychophysics*, *15*(2), 285–290.

Prieto, P., & Esteve-Gibert, N. (2018). *The development of prosody in first language acquisition*. Amsterdam: John Benjamins.

Sansavini, A., Bertoncini, J., & Giovanelli, G. (1997). Newborns discriminate the rhythm of multisyllabic stressed words. *Developmental Psychology*, *33*(1), 3–11.

Sazaki, T., Nakajima, Y., & Hoopen, G. T. (1998). Categorical rhythm perception as a result of assimilation in time-shrinking. *Music Perception*, *16*(2), 201–222.

Siegel, J. A., & Siegel, W. (1977). Categorical perception of tonal intervals: Musicians can't tell *sharp* from *flat*. *Perception & Psychophysics*, *21*(5), 399–407.

Shattuck-Hufnagel, S., & Turk, A. (1998). The domain of phrase-final lengthening in English. *Journal of The Acoustical Society of America*, *103*, 2889.

Singh, L., Nestor, S., Parikh, C., & Yull, A. (2009). Influences of Infant-Directed Speech on Early Word Recognition. *Infancy: The Official Journal of the International Society on Infant Studies*, *14*(6), 654–666.

Skoruppa, K., Pons, F., Christophe, A., Bosch, L., Dupoux, E., Sebastián-Gallés, N., Limissuri, R. A., & Peperkamp, S. (2009). Language-specific stress perception by 9-month-old French and Spanish infants. *Developmental science*, *12*(6), 914–919.

Slevc, L. R., & Miyake, A. (2006). Individual differences in second-language proficiency. *Psychological Science*, *17*, 675–681.

Spinelli, M., Fasolo, M., & Mesman, J. (2017). Does prosody make the difference? A meta-analysis on relations between prosodic aspects of infant-directed speech and infant outcomes. *Developmental Review*, *44*, 1–18.

Streeter, L. A. (1976). Language perception of 2-month-old infants shows effects of both innate mechanisms and experience. *Nature*, *259*(5538), 39–41.

Thiessen, E. D., Hill, E. A., & Saffran, J. R. (2005). Infant-Directed Speech Facilitates Word Segmentation. *Infancy: the official journal of the International Society on Infant Studies*, *7*(1), 53–71.

Trehub, S. E., Endman, M. W., & Thorpe, L. A. (1990). Infants' perception of timbre: Classification of complex tones by spectral structure. *Journal of Experimental Child Psychology*, *49*(2), 300–313.

Trehub, S. E., & Thorpe, L. A. (1989). Infants' perception of rhythm: Categorization of auditory sequences by temporal structure. *Canadian Journal of Psychology*, *43*(2), 217–229.

Trehub, S. E., & Thorpe, L. A. (1989). Infants' perception of rhythm: Categorization of auditory sequences by temporal structure. *Canadian Journal of Psychology*, *43*, 217–229.

Trehub, S. E., Thorpe, L. A., & Morrongiello, B. A. (1987). Organizational processes in infants' perception of auditory patterns. *Child Development*, *58*, 741–749.

Tsao, F. M., & Liu, H. M. (2020). Lexical-tonal perception development in infancy. In F., Tsao, P., Li (Eds.), *Speech perception, production and acquisition. Chinese language learning sciences*. Springer. https://doi.org/10.1007/978-981-15-7606-5_10

van Ommen, S., Boll-Avetisyan, N., Larraza, S., Wellmann, C., Bijeljac-Babic, R., Höhle, B., & Nazzi, T. (2020). Language-specific prosodic acquisition: A comparison of phrase boundary perception by French- and German-learning infants. *Journal of Memory and Language*, *112*, Article 104108.

Weisleder, A., & Fernald, A. (2013). Talking to children matters. *Psychological Science*, *24*(11), 2143–2152.

Wetzel, N., Buttelmann, D., Schieler, A., & Widmann, A. (2016). Infant and adult pupil dilation in response to unexpected sounds. *Developmental Psychobiology*, *58*(3), 382–392.

Wong, P. C. M., Ciocca, V., Chan, A. H. D., Ha, L. Y. Y., Tan, L.-H., & Peretz, I. (2012). Effects of culture on musical pitch perception. *PLoS ONE*, *7*(4), e33424.

Wu, H., Ma, X., Zhang, L., Liu, Y., Zhang, Y., & Shu, H. (2015). Musical experience modulates categorical perception of lexical tones in native Chinese speakers. *Frontiers in Psychology*, *6*, 436.

Yang, J., Kanazawa, S., Yamaguchi, M. K., & Kuriki, I. (2016). *Cortical response to categorical color perception in infants investigated by near-infrared spectroscopy. Proceedings of the National Academy of Sciences of the United States of America*, 113(9), 2370–2375.

Zhao, T. C., & Kuhl, P. K. (2016). Musical intervention enhances infants' neural processing of temporal structure in music and speech. *PNAS*, *113*(19), 5212–5217.

Zuk, J., Ozernov-Palchik, O., Kim, H., Lakshminarayanan, K., Gabrieli, J. D. E., Tallal, P., & Gaab, N. (2013). Enhanced syllable discrimination thresholds in musicians. *PLOS ONE*, *8*(12), e80546.

SUPPORTING INFORMATION

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

How to cite this article: Langus, A., Boll-Avetisyan, N., van Ommen, S., & Nazzi, T. (2023). Music and language in the crib: Early cross-domain effects of experience on categorical perception of prominence in spoken language. *Developmental Science*, e13383. https://doi.org/10.1111/desc.13383