Received: 26 May 2020 | Accepted: 2 February 2021

DOI: 10.1111/infa.12395

INFANCY THE OFFICIAL JOURNAL OF THE INTERNATIONAL CONGRESS OF INFANT STUDIES WILEY

RESEARCH ARTICLE

The development of infants' sensitivity to native versus non-native rhythm

Nikola Paillereau[1,2] | Václav Jonáš Podlipský[3] | Filip Smolík[1,4] | Šárka Šimáčková[3] | Kateřina Chládková[1,5]

[1]Institute of Psychology, Czech Academy of Sciences, Praha, Czech Republic

[2]Institute of Phonetics, Faculty of Arts, Charles University, Praha, Czech Republic

[3]Department of English and American Studies, Faculty of Arts, Palacký University, Olomouc, Czech Republic

[4]Department of Linguistics, Faculty of Arts, Charles University, Praha, Czech Republic

[5]Institute of Czech Language and Theory of Communication, Faculty of Arts, Charles University, Praha, Czech Republic

Correspondence
Nikola Paillereau, Institute of Psychology, Czech Academy of Sciences, Hybernská 8, 110 00 Praha, Czech Republic.
Email: paillereau@praha.psu.cas.cz

Funding information
Univerzita Karlova, Grant/Award Number: Primus/17/HUM/19; Grantová Agentura České Republiky, Grant/Award Number: 18-01799S

Abstract

Speech rhythm is considered one of the first windows into the native language, and the taxonomy of rhythm classes is commonly used to explain early language discrimination. Relying on formal rhythm classification is problematic for two reasons. First, it is not known to which extent infants' sensitivity to language variation is attributable to rhythm alone, and second, it is not known how infants discriminate languages not classified in any of the putative rhythm classes. Employing a central-fixation preference paradigm with natural stimuli, this study tested whether infants differentially attend to native versus nonnative varieties that differ only in temporal rhythm cues, and both of which are rhythmically unclassified. An analysis of total looking time did not detect any rhythm preferences at any age. First-look duration, arguably more closely reflecting infants' underlying perceptual sensitivities, indicated age-specific preferences for native versus non-native rhythm: 4-month-olds seemed to prefer the native-, and 6-month-olds the non-native language-variety. These findings suggest that infants indeed acquire native rhythm cues rather early, by the 4th month, supporting the theory that rhythm can bootstrap further language development. Our data on infants' processing of rhythmically unclassified languages suggest that formal rhythm classification does not determine infants' ability to discriminate language varieties.

KEYWORDS
early language development, infants, language preferences, speech rhythm, temporal cues, Czech

1 | INTRODUCTION

There is evidence of newborn infants' ability to discriminate native and non-native speech and of their preference of the former over the latter (Mehler et al., 1988; Moon et al., 1993, Byers-Heinlein et al., 2010). Early language discrimination has generally been attributed to newborns' innate sensitivity to three rhythm classes: stress-timed, syllable-timed, and mora-timed (Nazzi et al., 2000; but also Abboub et al., 2016; Gervain, 2018; May et al., 2017; Molnar et al., 2014; Mueller et al., 2015; Ramus et al., 2000; Thorson, 2018; Werker & Gervain, 2013; Werker & Tees, 2005). Using languages prototypically representing the 3 rhythm classes, newborns were indeed shown to discriminate languages from across but not from within the rhythm classes (Mehler & Christophe, 1995; Ramus et al., 2000). Within-class discrimination was attested for older infants, of about 4 months of age, having sufficient experience with native speech rhythm, which they could distinguish from an unfamiliar but "similar"[1] rhythm (Bosch & Sebastián-Gallés, 1997; Molnar et al., 2014; Nazzi et al., 2000).

Speech rhythm generally, and specifically the hypothesis positing the existence of the abovementioned rhythm classes, has thus in the literature been considered central to infants' ability to discriminate languages. However, the primacy given to speech rhythm and to the rhythm classes in studies on early language discrimination and language preferences is to date not unequivocally supported by empirical data.[2]

Firstly, the importance of formal rhythmic classification advanced by previous studies is biased by the choice of stimuli because only languages prototypical of the respective rhythm classes were used (except for Catalan used in Bosch & Sebastián-Gallés, 1997, whose status is questionable, for example, Grabe & Low, 2002). The question remains how infants respond to differences between languages which are not straightforwardly classifiable in terms of the rhythm taxonomy. For infants to discriminate between two languages, do these languages really need to fall into distinct rhythm classes or is it sufficient if they display salient differences in the acoustic correlates of rhythm? The present study tests the hypothesis that infants' early sensitivity to language variation is predicted by the acoustics of language rhythm, rather than restricted to the traditional three-way rhythm classification. To overcome the limitations by the formal rhythm taxonomy, the present study focuses on Czech, which is a rhythmically unclassified language (Dankovičová & Dellwo, 2007).

Secondly, besides the lack of infant perception data on rhythmically unclassifiable languages, even within the broad literature dealing with languages that are prototypes of the rhythm classes, the exclusive role of rhythm in infants' perception of speech is not evidenced conclusively. Although the early

[1]The term "rhythmically similar" is commonly used in infant literature to describe languages that belong to the same rhythm class within the traditional classification.

[2]Some previous studies tested infants' sensitivity to local rhythm phenomena with isolated CVCV strings, and showed that from birth on, infants are sensitive to the local stress changes, that is, to the difference between a trochaic and iambic pattern of (nonce) words (Jusczyk & Thompson, 1978; Sansavini et al., 1997). It is however unclear whether infants actually employ this sensitivity when listening to continuous spoken utterances produced in different languages.

TABLE 1 Studies using a preferential design on language (variety) preferences in infants aged 0 to 9 months

References	Procedure/Measure	Preference/age (in months) 0	3	4	5	6	7	8	9	Notes
Bosch and Sebastián-Gallés (1997)	Visual Orientation Procedure			F						
	Training phase (language not used at test)			F						LPF stimuli
	Mean reaction time			N						Bilingual infants
Bosch (1998)	Modified Head-Turn Preference Procedure Familiarization phase Mean orientation time			N						
Nazzi et al. (2000)	Modified Head-Turn Preference Procedure				N					
	Familiarization phase				N					Unfamiliar lgs., different classes
	Mean orientation time				X					Unfamiliar lgs., same class
Moon et al. (1993)	High-Amplitude Sucking Preference Procedure NO habituation Mean number of sucks	F								
Byers-Heinlein et al. (2010)	High-Amplitude Sucking Preference Procedure	F								LPF stimuli
	NO habituation Mean number of sucks	X								LPF stimuli, bilingual infants
Butler et al. (2011)	Modified Head-Turn Preference Procedure				F		N			
	Familiarization phase Mean orientation time				X					Both lgs. unfamiliar

(Continues)

TABLE 1 (Continued)

References	Procedure/Measure	Preference/age (in months)								Notes
		0	3	4	5	6	7	8	9	
Kitamura et al. (2006)	Visual fixation preference NO habituation Mean looking time		F			X N				Both lgs. familiar
Kitamura et al. (2013)	Visual fixation preference NO habituation Mean looking time		F			X F			X	Both lgs. familiar
Diehl et al. (2006)	Visual fixation preference NO habituation Mean looking time					N X		X		LPF stimuli

Note: Studies with an infant-controlled paradigm are in grey. The Notes column marks deviations from the standard design, which includes monolingual infants, non-filtered speech, and a paradigm with one familiar and one unfamiliar language.

Abbreviations: F, familiarity preference; LPF, low-pass filtered; N, novelty preference; X, no preference.

language discrimination, found across studies and language backgrounds, has generally been ascribed to infants' sensitivity to speech rhythm, to date there are insufficient data to determine whether, at what age, and to what extent, infants are sensitive to rhythm variation alone. To the best of our knowledge, only one infant study so far, Ramus (2002), tested the role of suprasegmental features of speech, that is, rhythm and intonation, separately (using delexicalized utterances as stimuli, in which all the different vowels and consonants are replaced by always the same vowel or consonant, preserving original segment durations). Ramus' (2002) results on rhythm alone, however, are inconclusive: within that study, the experiment for which the entire 2-min test phase was analyzed provided no evidence of discrimination, unlike a post-hoc analysis of the 1st minute only (showing discrimination). Other infant studies interested in rhythm used speech materials containing not only temporal rhythm cues, but also phrase-level intonation patterns. These studies used either synthesized stimuli (Ramus et al., 2000) or low-pass filtered stimuli (Bosch & Sebastián-Gallés, 1997; Diehl et al., 2006; Jusczyk et al., 1993; Molnar et al., 2014; Nazzi et al., 1998) preserving suprasegmental information in the speech signal while removing most (though not all) segmental information (i.e., characteristics of different vowels and consonants). Other studies, interested in the role of rhythm or rhythm classes, used even cues to language- or variety-specific phonemic categories (Bosch & Sebastián-Gallés, 2001; Butler et al., 2011; Diehl et al., 2006; Kitamura et al., 2006, 2013; Nazzi et al., 2000; Phan & Houston, 2009).[3] The present study addresses this problem and to test the role of rhythm alone it employs (naturally produced) stimuli that differ in the temporal relations between strong and weak syllables but not in other respects.

As explained above, the present study aims to determine whether the acoustic rhythmical properties of speech and not the classification of a language into a formal rhythm class is what drives infants' sensitivity to language variation. Let us now explain why determining this is crucial for our understanding of early language development. It has been argued that infants' differential preferences for native versus non-native rhythm patterns do not merely mark the ability to discriminate but shows that the native rhythm patterns have been acquired (Keij, 2017). Knowing when native rhythm is acquired is important in order to understand further development of speech and language. If rhythm preferences indeed emerge soon after birth, it is possible that the knowledge of native rhythm patterns giving rise to such preferences acts as a bootstrapping mechanism for further language learning (as argued by, e.g., Gervain, 2018). For instance, having mastered native speech rhythm could facilitate syntactic learning (Gleitman & Wanner, 1982) and word segmentation (Nazzi et al., 2006). The knowledge of native rhythm could help infants bootstrap into many of the other types of knowledge, provided that rhythm as such is indeed acquired early in development as has been argued (or even simply assumed) previously.

Table 1 lists studies that examined whether young infants manifest differential attention to native- and non-native rhythm when listening to continuous speech displaying a rich variety of cues (note that only three preference studies used low-pass filtered stimuli). The table shows that when using a non-infant controlled familiarization-preference paradigm, in which infants listen to one of the languages (or language varieties) during a familiarization phase, they prefer the other language (variety) during the test, that is, they manifest a novelty preference (Bosch, 1998; Nazzi et al., 2000). When testing spontaneous (pre-experimental) preferences without a familiarization phase to one of the languages under test, infants aged 0–5 months pay more attention to their native language over a non-native one, that is, they manifest a familiarity preference (Butler et al., 2011; Byers-Heinlein et al., 2010; Diehl

[3]Diehl et al. (2006) used both natural stimuli and low-pass filtered stimuli.

et al., 2006; Kitamura et al., 2006; Kitamura et al., 2013; Moon et al., 1993), except for bilingual infants who do not manifest preference at birth (Byers-Heinlein et al., 2010) or manifest a novelty preference at 4 months (Bosch & Sebastián-Gallés, 1997). By contrast, infants aged 6 or more months either pay more attention to the non-native language over their native one, that is, they manifest a novelty preference[4], or they do not exhibit differential attention. Finally, if the tested languages (varieties) are both non-native and unfamiliar, infants do not seem to have any preferences (Butler et al., 2011; Nazzi et al., 2000), which supports the idea that if none of the languages is familiar, infants will attend similarly to both while if one *is* familiar—and acquired—infants will exhibit preferences (see Keij, 2017).

The main cue contributing to the perception of rhythm is most likely duration, as reflected by many of the widely used rhythm metrics. Among the different rhythm metrics operating with segmental duration, VarcoV and %V[5], that represent the variability of vowel duration, and the proportion of vowels within an utterance, seem to reflect listeners' behavior most reliably, especially when also dialectal differences are taken into account (White & Mattys, 2007a). Using monotonized and delexicalized utterances, White et al. (2012) showed that on the basis of durational contrasts between stressed and unstressed syllables, which determine speech rhythm, English adults distinguish better between regional varieties of their L1 than between English and Dutch and that rhythm differences can be larger within a single rhythm class than across two rhythm classes. Language (variety) discrimination was thus better predicted by the magnitude of the durational contrasts rather than by formal classification in terms of the putative rhythm classes. It is thus plausible that infants' perception of language differences, too, is driven by sensitivity to the acoustic durational contrasts.

To test whether infants attend differently to utterances varying in the durational rhythm contrast, that is, in the temporal patterning of stressed and unstressed syllables, we employed a novel stimulus design. We devised two rhythm varieties of the infants' native language (Czech). Native female speakers produced well-formed sentences with low-frequency words with a rhythm typical of the native language and with an atypical, that is, non-native rhythm. The two varieties were acoustically rhythmically distinct from one another (see Figure 2 in the Methods) but neither of them is classifiable in terms of the traditional rhythm taxonomy. We assessed infants' perceptual preferences for the two rhythmic varieties across early development, between 3.5 and 10.5 months of age.

The two language varieties that we test (Czech and Czech with atypical rhythm) are rhythmically unclassified but differ largely in acoustic measures of rhythm (VarcoV and %V). We predict that infants will discriminate between them and manifest preferences at the earliest tested age, that is, at 3.5–4.5 months. Next, in line with the age-dependent diminishing of preferential attention or discrimination shown previously (Table 1), we predict that by the oldest ages tested, that is, 9.5–10.5 months, the infants will no longer demonstrate preferential listening to one rhythmic variety over the other (probably as an effect of having turned their attention to other linguistic features such as segmental or lexical information). As for the direction of preferences in the younger infants, we did not formulate any a priori predictions; however, from Table 1 it appears that the youngest (monolingual) infants in general prefer native or familiar materials, while slightly older ones prefer nonnative or unfamiliar ones (before the preferences disappear entirely at an even older age).

[4]Relatedly, a recent meta-analysis by Gasparini et al. (2020) detected that increasing infant age predicts a novelty preference when the stimuli are languages from the same rhythm class.

[5]VarcoV is a rate-normalized measure calculating standard deviation of vocalic interval duration divided by mean vocalic duration (and multiplied by 100), and V% represents the percentage of vocalic intervals' duration within an utterance.

2 | METHODS

2.1 | Participants

Eighty-nine monolingual Czech-learning infants, aged 3.5–10.5 months, participated. Ten additional infants were excluded due to fussiness ($n = 6$), technical error ($n = 1$), or bilingualism ($n = 3$). The 89 infants completed at least four out of the ten trials (i.e., at least two native and two non-native trials). The count and ages per group were: 23 four-month-olds (nine girls, mean age 123 days, range 106–142), 18 six-month-olds (11 girls, mean age 184 days, range 169–208), 27 eight-month-olds (13 girls, mean age 242 days, range 227–256), and 21 ten-month-olds (11 girls, mean age 304 days, range 290–316). All were full-term, healthy, and raised by Czech-speaking parents.

Recruitment went through social networks and family centers. The experiment followed the standards of the Declaration of Helsinki was approved by the ethical committee of the Institute of Psychology, Czech Academy of Sciences, and administered after parental informed consent. The families received a voucher for participation.

2.2 | Stimuli

Auditory stimuli were created as follows. First, a Czech female speaker produced 30 nonsensical but grammatically well-formed sentences containing only low-frequency words in infant-directed speech (IDS), for example, *Klopa rohovky se vzdouvá* "The lapel of the cornea is billowing." These naturally produced sentences were then also edited to create versions with non-native rhythm: using overlap-add resynthesis in Praat (Boersma & Weenink, 2020), the (foot-initial) stressed syllables were lengthened at the expense of the following unstressed syllable(s) in each foot, preserving original foot (and sentence) durations. The lengthening of each stressed syllable varied linearly between 0% and 120% depending on the stressed-syllable/foot duration ratio. The 30 natural as well as the 30 edited productions of the model speaker were subsequently imitated by three other Czech women, yielding naturally produced renditions with native segments but with native and non-native rhythmic patterns. Four members of the author team who are phoneticians (all native speakers of Czech) made sure that the final "native" and "atypical" stimuli differed in the target feature, that is, the timing of stressed and unstressed syllables, rather than the intonation contours, global tempo, fluency, naturalness, the degree of infant-directed-speech modulations, and the properties of the vowels and consonants contained. Figure 1 shows spectrograms of an example sentence with native and with non-native timing.

From each of the three speakers, we selected 15 native imitations and 15 non-native rhythm imitations of different sentences. In the final stimulus set, thus, one speaker never uttered the same sentence in both rhythms. Furthermore, we ensured that the final stimulus set contained all the 30 different sentence identities, comparably distributed across the three speakers and the two language varieties.

The final set of 90 sentences (3 speakers * [15 native + 15 non-native]) was used to create 10 stimulus trials, each containing nine different sentences, three from each speaker. Half of the trials, that is, five were spoken in native rhythm, the other five in non-native rhythm. Each native trial was matched with a nonnative trial, as far as possible, in the identities of the nine sentences it contained (effectively resulting in five native-nonnative trials pairs). Trials lasted between 21.2 and 25.2 s.

We compared the durations, mean F0 and intensity of vocalic intervals, and within-vowel F0 and intensity ranges, in stressed and unstressed syllables in the two language varieties with a factorial analysis of variance. For two variables, significant interactions between stress and rhythm type were revealed ($F > 28$, $p < .001$): vowel duration (as expected) and F0 range within a vowel (corresponding to vowel-internal

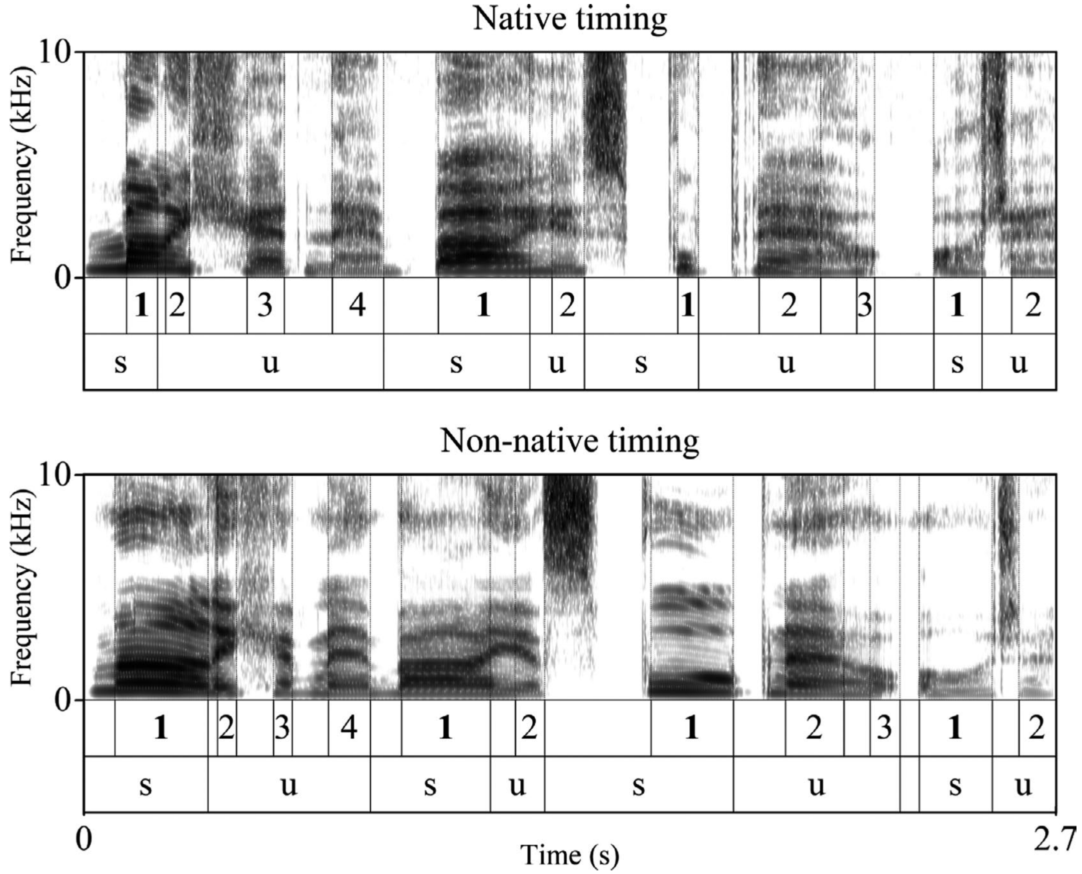

FIGURE 1 Spectrograms of the phrase *Malicherné báje zpupného oře*. "Trifling myths of a mutinous steed.", recorded with native Czech timing and (by another speaker) with non-native timing whereby word-initial stressed syllables were produced as prolonged. The stressed (word-initial) syllables are marked as "s", the associated stressed vowels as boldface "1"; unstressed syllables in each foot as "u" and unstressed vowels within each foot as "2," "3," or "4." Note the lengthened stressed and shortened unstressed syllables in the non-native utterance

change of pitch); mean values for these variables are given in Table 2. The measured differences indicate that the native and non-native rhythm differed markedly in durational patterns and somewhat also in F0 change within vowels, which slightly enhanced the prominence of stressed syllables and it was probably a by-product of vowel lengthening. Additionally, we measured the VarcoV and %V, metrics of rhythm (Dellwo, 2006; Ramus et al., 1999; White & Mattys, 2007). Figure 2 compares these metrics measured for the natural Czech stimuli and atypical-rhythm stimuli with those published previously for 15 other languages and language varieties. One-way analyses of variance showed that both VarcoV and %V were significantly higher in the atypical than in the natural Czech stimuli ($F[1, 88] > 34$, $p < .001$).

2.3 | Procedure

To test infants' perceptual acquisition of speech rhythm, we employed a central-fixation preference paradigm without familiarization. Arguably, such paradigm has an advantage over a habituation–dishabituation paradigm in that it tests not only (acoustically driven) discrimination abilities but also the

TABLE 2 Vowel duration and F0 range within stressed and unstressed vowels in the native- and non-native-accented sentences

	Duration (ms)		F0 range (semitones re 100 Hz)	
	Mean	95% conf. int.	Mean	95% conf. int.
Native				
Stressed V	103.3	95.0–111.5	2.10	1.74–2.46
Unstressed V	100.7	94.5–106.9	2.08	1.81–2.35
Non-native				
Stressed V	208.2	200.0–216.4	3.60	3.25–3.95
Unstressed V	84.3	78.0–90.6	1.87	1.59–2.14

degree of pre-experimental acquisition of previously encountered native speech patterns (Houston-Price & Nakai, 2004). Some preference procedures, such as (the modifications of) a head-turn preference paradigm (HPP), require a familiarization phase, in which infants get familiarized with the location of the laterally placed sound (Bosch, 1998; Bosch & Sebastián-Gallés, 1997; Butler et al., 2011; Nazzi et al., 2000). The familiarization phase often uses one of the languages (variants) that occurs at test, leading to results that are not informative about spontaneous preferences but only about discrimination. In contrast, visual fixation paradigms with centrally presented audio-visual stimulation (CF, used in the present study, and in Diehl et al., 2006; Kitamura et al., 2006, 2013), do not require familiarization. Preference procedures without a familiarization thus straightforwardly reflect the naturalistic, pre-experimental preferences and thus also acquisition (Cristia et al., 2012; Houston-Price & Nakai, 2004). It should be noted here that an absence of preference should not be interpreted as an absence of discrimination; it might be that infants discriminate the two types of stimuli but find them equally perceptually attractive.

In the present experiment, infants were seated in a sound-treated booth on their parent's lap in front of a 19-in screen, 1 m away. The screen displayed a multi-color static checkerboard during test trials and a rotating color wheel between trials, serving as an attention-getter (with the sound of chimes). Auditory stimuli were presented through two loudspeakers located (behind a curtain) on each side of the monitor at a comfortable listening level of approximately 65 dB SPL. Test trials were initiated when the infant looked for two consecutive seconds to the immediately preceding attention-getter. Test trials were played at full length (21.2–25.2 s) irrespective of the infant's behavior. The full experiment contained 10 trials, that is, 5 of each language variety that alternated in pseudorandom order. Infants were included if they completed at least four out of the 10 trials. The experimenter was blind to the order of trials.

A video camera placed frontally recorded the infant and registered trial start and end markers. The parent was instructed not to interact with their child and listened to masking music. The experiment ran on PyHab (Kominsky, 2019) within PsychoPy (Peirce et al., 2019). Outside of the sound-treated testing booth, the experimenter monitored infant behavior and held a button whenever the infant was looking at the screen.

2.4 | Data coding, measures, and statistical analyses

Looks were offline-coded using ANVIL (Kipp, 2014). The coder, blind to the stimuli as she only received a silent video output, marked all looks to the screen. A separate sound output containing beep-marked trial boundaries was subsequently integrated with the looking-data annotations.

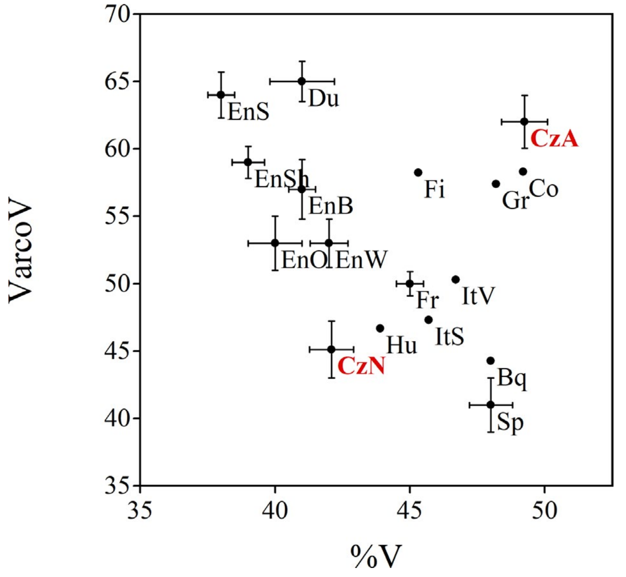

FIGURE 2 VarcoV by %V plot showing scores for the natural (CzN) and atypical (CzA) Czech stimuli used in the present study, in comparison with those for Du = Dutch, Fr = French and Sp = Spanish (White & Mattys, 2007), for several varieties of English: EnS = Standard Southern British English, EnSh = Shetland English, EnB = Bristol English, EnO = Orkney English, and EnW = Welsh Valleys English (White & Mattys, 2007a), for ItV = Venetian Italian, and ItS = Sicilian Italian (White et al., 2009), for Hu = Hungarian, Fi = Finnish (data reported in White et al., 2012), for Co = Corean, Gr = Greek (Arvaniti, 2012), and for Bq = Basque (Molnar et al., 2014). Dots show mean values, whiskers for the previously published data represent 1 *SD*, whiskers for the present data on Czech represent 1 *SEM*

For statistical analyses, we extracted two measures, namely, total looking time normalized by trial length, and first-look duration, that is, the interval starting when an infant first looked at the screen up to the first look away. The duration of the first look that we employ here is an alternative, and in the present paradigm probably more reliable, measure to trace infants' preferences (Aldridge et al., 1999; Lewis et al., 1966).

Recall that the testing paradigm used here was not contingent on infant looking behavior, that is, stimuli were presented in full irrespective of whether the infant was looking at the screen or not. The reason why we chose to present all stimuli in full was to ensure that all participants had the same amount of exposure to the native and nonnative rhythm, that is, that particularly the nonnative condition remained comparably nonnative across all participants. However, a paradigm that is not contingent on infant behavior may not be optimal in revealing infants' actual stimulus processing/preferences. As pointed out by Sundara et al. (2018) in the non-contingent paradigm, infants may not grasp the coupling between stimulus presentation and looking behavior: a lack of looking at the screen does not imply infants' lack of attention or lack of preference to an auditory stimulus. Inversely, we suggest that infants may keep looking or re-orient towards the stimulation, even if they do not find it interesting, simply because the stimulus is kept on. That is, infants who are being tested in a dark environment with the only source of light coming from a monitor may not necessarily express their preferences for an auditory stimulus by looking towards its source. For this reason, it has been suggested that rather than a total time of all recurrent looks towards auditory-visual stimulation within a trial, it is the pattern of looking behavior which elicits preferences: a long-lasting look followed by several short looks might reflect preferences better than many short looks even if the total look for both patterns is the same (Lewis et al., 1966).

To make the story even more complicated, such failures to figure out the coupling between stimulus preferences and looking behavior might, possibly, be more frequent or more variable in younger than in older infants, because younger infants may have greater difficulty with disengagement of visual attention (Hood & Atkinson, 1993). As seen above, considering an infants' total looking time during trials in a non-contingent experimental paradigm may not always reliably reflect the infants' attentional processing or underlying preferences for the stimuli at hand. This disadvantage of non-contingent paradigm could be at least partially compensated by considering only the duration of the first look in the individual trials. That is, instead of stopping the stimulus once the infant looks away (which is what a contingent presentation paradigm would do), one can stop measuring the looking time once the infant looks away from the screen for the first time (which is what taking the duration of the first look does). We thus believe that the first look to the screen more veridically reflects an infant's true attentional/underlying preferences than do any subsequent looks. Note that we are primarily interested in how the first look duration differs when listening to native versus non-native rhythm within each age group, which means that the differential ease with which infants engage and disengage at different ages is not counter-indicative for the use of the first look measure.

The individual trial looking times were analyzed with linear mixed models, using *lme4*and *lmerTest* (Bates et al., 2015; Kuznetsova et al., 2017) in R (R Core Team, 2019). The dependent variable was first look duration in seconds. The fixed effects were trial type (non-native vs. native, coded −1 vs. +1), infant age in months (with three contrasts comparing each younger to each closest older group; coding the 4-, 6-, 8-, 10-month as −1, +1, 0, 0 for the first, 0, −1, +1, 0 for the second, and 0, 0, −1, +1 for the third contrast), their interaction, and trial length (mean-centered). The random-effects structure contained per-subject and per-item (i.e., the five different trial pairs) intercepts and slopes for trial type.

3 | RESULTS

Figures 3 and 4 plot the measured total looking time (normalized for trial length) and first look durations, respectively, and Tables 3 and 4 give the model outputs for the two measures.

For total looking time (Figure 3 and Table 3), there was a significant intercept: the average looking time was 13.652 s. Two age contrasts yielded significant main effects showing that 8-month-olds looked shorter than the 6-month olds, by on average 2.6 s, and that the 10-month-olds looked shorter than the 8-month-olds, by on average 2.9 s. A significant effect of trial number shows that infants' looking time decreased with increasing trial number.

For first look duration (Figure 4 and Table 4), there was also a significant intercept: the average first look duration was 5.958 s. A significant effect of trial number revealed that first look duration decreased with increasing trial number. There was an interaction of trial type and the first age contrast, suggesting that 4-month-olds and 6-month-olds differed in their preferential looking to native versus non-native trials.

To unpack the significant two-way interaction, we compared the estimated marginal means and 90% confidence intervals per trial type per age group, using *ggeffects* (Lüdecke, 2018); the comparisons are visualized in Figure 5 and summarized in Table 5.

The pairwise comparisons show that 4-month-olds had longer first look to the native than to the non-native trials by on average 0.95 s (native mean = 7.048 s, CI = 6.277–7.819; non-native mean = 6.097, CI = 5.432–6.762), while the 6-month-olds had longer first look to the non-native than to the native trials by on average 1.39 s (non-native mean = 6.280, CI = 5.101–7.460; native mean = 4.890, CI = 3.578–6.202). The comparisons for the 8- and 10-month-olds did not turn out meaningful as the confidence intervals of their first look duration in native and non-native trials largely overlapped.

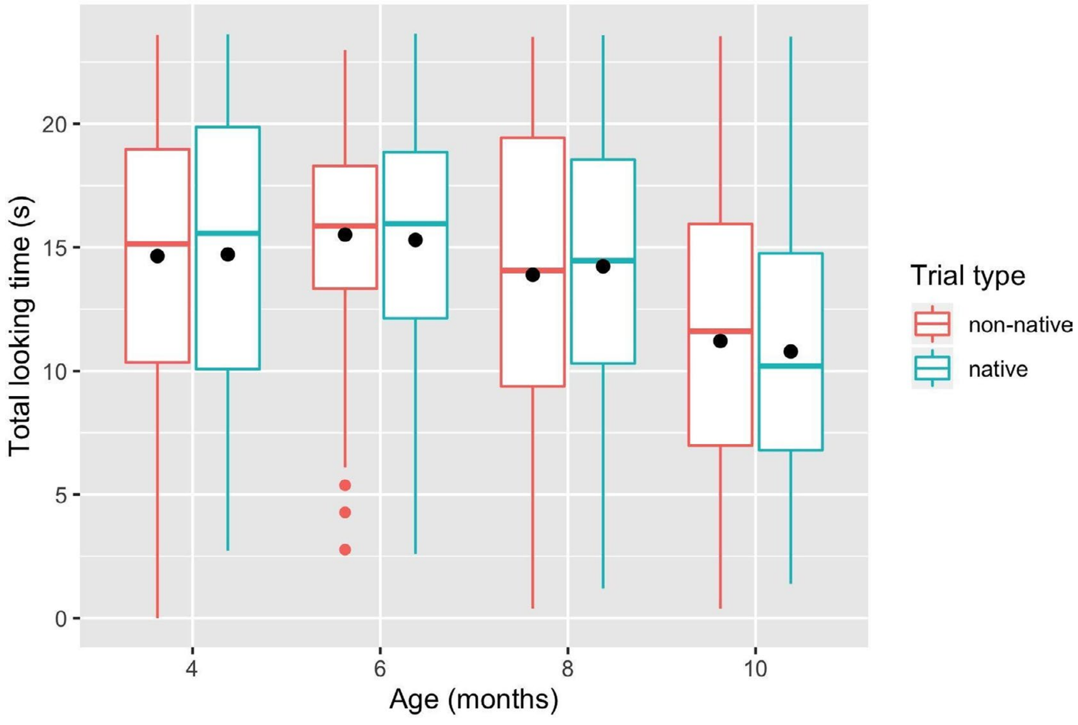

FIGURE 3 Plot of the total looking time data; because individual trials varied in length, the measured total looking time was normalized for mean trial length (i.e., the proportion of total looking in each trial was multiplied by average trial length 23.7 s). Black dots show means, horizontal lines within boxes show medians, boxes show the interquartile (IQ) range, whiskers the 1.5 * IQ range, dots are potential outliers

To evaluate whether our sample sizes were adequate, we performed simulated power analyses using the simr R package (Green & MacLeod, 2016), estimating across 1000 simulations. For the total looking time model, the post-hoc power for the main effect of trial type was 9% (which is not surprising, as the small effect and its high *p*-value indicate that the power was low). To achieve 80% power for an effect of the size that we observed, we would need to test about 1800 participants. A larger effect would be of more interest, so we used repeated simulations to estimate the size of the effect that would be associated with 80% power in 80 participants. The resulting difference between trial types was approximately 0.38 second of the normalized looking time, which is less than 3% of the variable's standard deviation. The power to detect interesting effects thus appears quite solid.

For the first look duration, the observed power for the main effect of trial type (experimental condition) was 8%. The number of participants to achieve 80% power for this effect was around 170. The post-hoc power for the significant interaction between trial type and the first age contrast, that is, for the difference between the 4- and 6-month-olds' data, was 58.6%. To achieve 80% power for this effect, the sample would have to comprise slightly more than 150 participants. The observed power of 58.6%[6] is lower than the usually desired standard of 80%.

[6]Although replicating/repeating previous methodological flaws does not make them less flawed, we would like to note that our relatively low power is quite in line with the power of infant studies asking comparable questions and employing comparable paradigms; for instance, in studies of infants' preferences for infant-directed speech the power is typically about 60% (Bergmann et al., 2018).

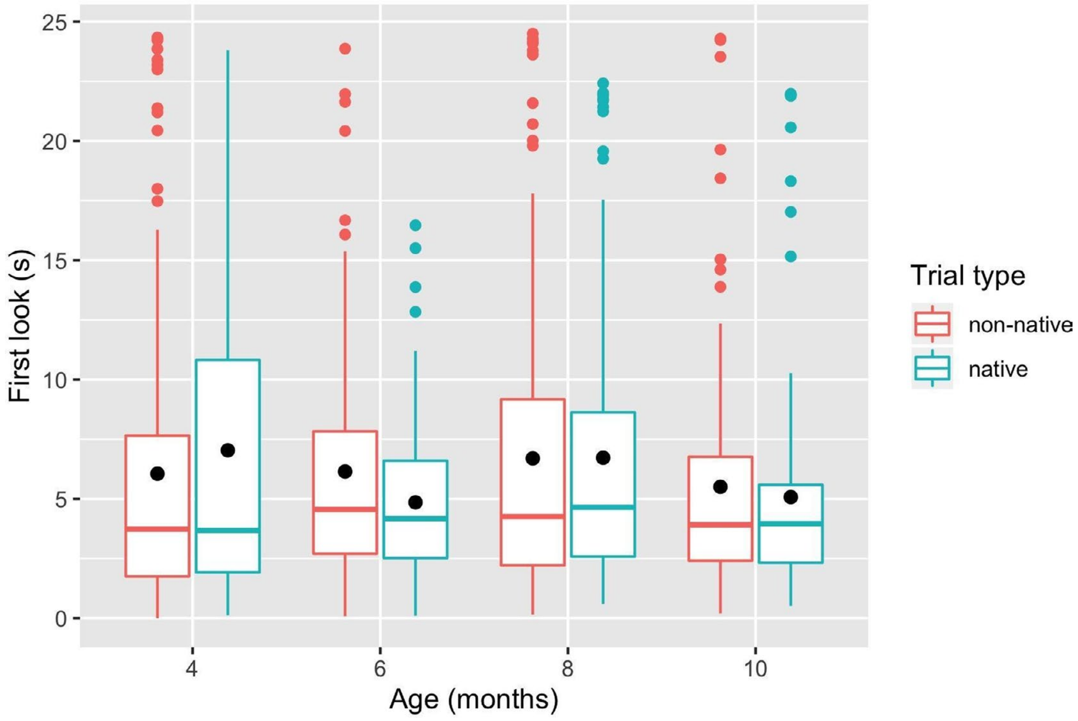

FIGURE 4 Plot of the first look duration data. Black dots show means, horizontal lines within boxes show medians, boxes show the interquartile (IQ) range, whiskers the 1.5 * IQ range, dots are potential outliers

TABLE 3 The fixed-effects output of the linear mixed model for total looking time

Predictor	Estimate	*SE*	*df*	*t*	*p*
Intercept	**13.652**	**0.439**	**59.944**	**31.101**	**<.001**
Trial Type (-nonnative +native)	−0.088	0.156	6.420	−0.561	.594
Age contrast 1 (−4 months +6 months)	−0.921	0.715	83.730	−1.288	.201
Age contrast 2 (−6 months +8 months)	**−2.571**	**0.843**	**84.532**	**−3.052**	**.003**
Age contrast 3 (−8 months +10 months)	**−2.906**	**0.741**	**85.686**	**−3.920**	**<.001**
Trial number (mean-centered)	**−0.688**	**0.058**	**7.682**	**−11.918**	**<.001**
Trial type * Age contrast 1	−0.088	0.217	680.790	−0.403	.687
Trial type * Age contrast 2	−0.024	0.260	684.916	−0.091	.928
Trial type * Age contrast 3	−0.223	0.233	686.854	−0.956	.339

Note: Significant effects are in bold.

4 | DISCUSSION

This study addressed the question of whether the well-documented early abilities of human infants to distinguish natural languages can be solely due to infants' sensitivity to the dynamics of syllable timing and prominence, that is, actual speech rhythm, and whether native rhythm is acquired early in life. Theories of language acquisition propose that prosody bootstraps further language learning (Gleitman & Wanner, 1982; Hirsh-Pasek et al., 1987; Jusczyk et al., 1992; Nazzi et al., 2006) and

TABLE 4 The fixed-effects output of the linear mixed model for first look duration

Predictor	Estimate	*SE*	*df*	*t*	*p*
Intercept	**5.958**	**0.404**	**14.719**	**14.730**	**<.001**
Trial type (−nonnative +native)	−0.136	0.181	10.507	−0.751	.469
Age contrast 1 (−4 months +6 months)	−0.498	0.578	81.935	−0.862	.391
Age contrast 2 (−6 months +8 months)	−0.010	0.684	83.565	−0.014	.989
Age contrast 3 (−8 months +10 months)	−0.779	0.604	85.883	−1.290	.201
Trial number (mean-centered)	**−0.514**	**0.090**	**4.177**	**−5.728**	**.004**
Trial type * Age contrast 1	**−0.611**	**0.281**	**679.865**	**−2.179**	**.030**
Trial type * Age contrast 2	−0.052	0.335	683.600	−0.155	.877
Trial type * Age contrast 3	−0.153	0.301	685.385	−0.507	.612

Note: Significant effects are in bold.

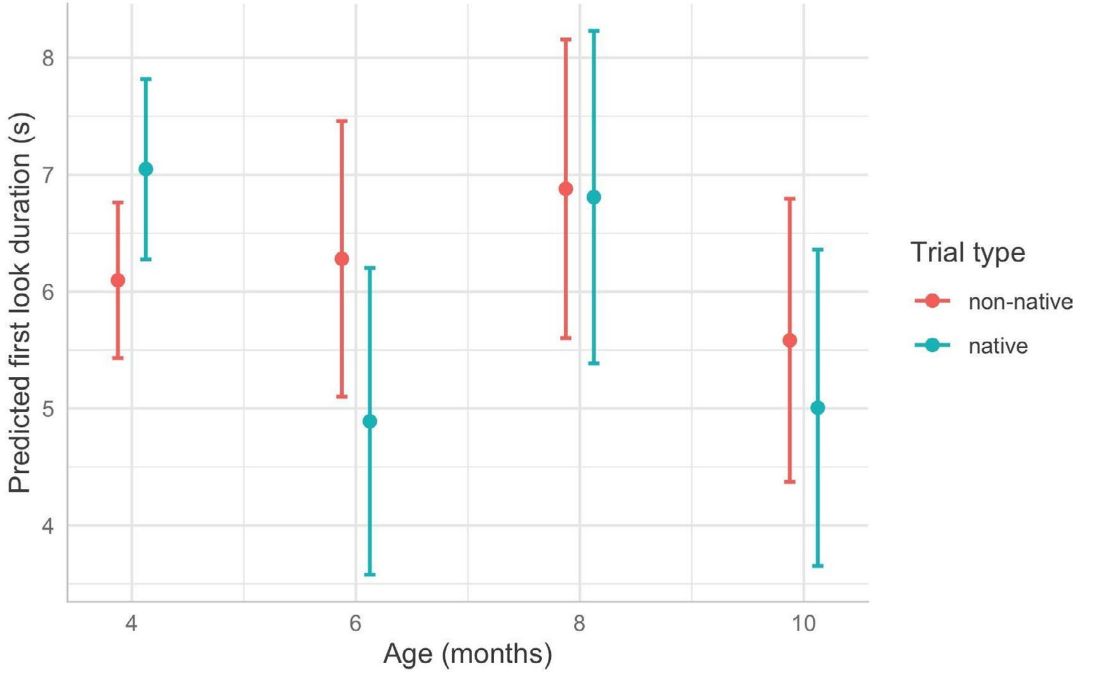

FIGURE 5 Model-predicted means of first look duration and 90% confidence intervals per age and trial type

seem to attribute a pivotal role to global rhythm in particular (Molnar et al., 2014; Nazzi et al., 2000, 2006) albeit without direct evidence. This study sought to determine whether the language-specific patterns of temporal relations in speech (i.e., rhythm) are indeed acquired early in development and could thus serve as a bootstrapping mechanism. We assessed infants' preferences for native versus non-native language varieties that differ in rhythm cues only. Infants aged 4, 6, 8, and 10 months acquiring Czech, a language that cannot be classified in terms of any of the three rhythm classes, were tested on their preferential looking to well-formed Czech sentences containing low-frequency lexical items, spoken in IDS mode, and produced naturally with native and non-native speech rhythm.

TABLE 5 Modelled marginal means and confidence intervals per age and trial type, and effect sizes (calculated as differences between the non-native and native means)

	Non-native rhythm		Native rhythm		
Age	**Mean**	**90% CI**	**Mean**	**90% CI**	**Simple effect size**
4 months	6.097	5.432–6.762	7.048	6.277–7.819	−0.951
6 months	6.280	5.101–7.460	4.890	3.578–6.202	1.390
8 months	6.879	5.602–8.157	6.809	5.385–8.232	0.070
10 months	5.583	4.370–6.796	5.006	3.652–6.360	0.577

We analyzed the typically employed measure of total looking time, and a less typical—but arguably more informative—measure of first look duration. No main or interaction effects involving rhythm were detected for total looking time. However, for first look duration, the analyses revealed developmentally conditioned rhythm preferences. Firstly, rhythm preferences were found in the two youngest groups, indicating that already at 4- and 6-months infants discriminate and have differential preferences for native versus non-native rhythm patterns. Interestingly, the direction of preference interchanged between 4 and 6 months. At 4 months, infants looked longer to native than to non-native trials, while the 6-month-olds looked longer to non-native than to native rhythm. The findings for first look thus indicate that native rhythm is acquired already within the first half of the infants' first year. Such acquisition of native language rhythm within the first half of the first year of life is in line with literature which argued that early-developed rhythm knowledge can bootstrap the acquisition of higher linguistic levels (Gervain, 2018; Nazzi et al., 2000). However, these studies did not unequivocally show that infants' early linguistic abilities are attributable to rhythm in particular. With carefully controlled but fully naturalistic stimulus material the present study demonstrates that the acoustic manifestation of rhythm alone is sufficient for young infants to discriminate languages.

Our experiment suggests that it is infants' sensitivity to the acoustic cues to rhythm, namely the temporal relations between strong and weak syllables, rather than the formal rhythm class distinctions, that drives the here-reported—as well as the earlier documented—infants' abilities to discriminate languages. Both language types compared in the present study do not lend themselves to straightforward classification in terms of the traditional rhythm classes, yet, the infants were able to discriminate between them as early as at 4 months of age. This finding highlights the role of the perceptual salience of the acoustic distance between different speech rhythm patterns in language discrimination and questions the role of the established rhythm classes in language development.

Let us now turn to the shift in the direction of preference between 4 and 6 months of age. Notoriously, directionality in infant behavioral paradigms is difficult to interpret. The predominant direction of preferences can vary across tasks, with, for example, a typical familiarity preference in word segmentation tasks, and novelty preference in rule learning tasks (Bergmann & Cristia, 2016; Bergmann et al., 2019). As reviewed in the Introduction and summarized in Table 1, the valence of preferences found in our study corresponds to that generally found across previous studies on language preferences, some of which used paradigms roughly comparable to ours. The interpretation of the directionality of the effects in each individual study is far from obvious (Rabagliati et al., 2019). For instance, two studies (Kitamura et al., 2006; Kitamura et al., 2013) using the same experimental design to test language-variety preferences at 6 months of age report a novelty and a familiarity preference, respectively. Kitamura et al. (2006) and Kitamura et al. (2013) argued that both directions indicate true preferences and proposed an explanation of the apparent discrepancy in terms of familiarity with the languages under test: whereas familiarity preference was found for a native versus a non-native unfamiliar language variety, novelty preference was found for a non-native, but familiar language variety. Our results do not align with that speculation and instead point towards a developmental shift from familiarity to novelty preference, warranting further empirical and/or meta-analytic investigations. While it is difficult to determine which preference valence in a single (insufficiently powered) study is genuine and which could possibly be a sign error (Bergmann et al., 2018; Gelman & Carlin, 2014), the age-related shift from familiarity to novelty preference was found for language pairs within the same formal rhythm class in a recent meta-analysis by Gasparini et al. (2020). However, such a result was not found for languages formally described as belonging to different rhythm classes and since our language variants are rhythmically unclassified (and our purpose is not to categorize them into discrete rhythm categories), the meta-analysis is neither supportive of nor inconsistent with our findings.

In sum, setting the directionality issue aside, the present study indicates that within the first half of their first year infants have acquired native temporal rhythm. This finding is based on one of two analyses carried out here, namely that of first look duration. A null result with respect to rhythm preferences was found in the other of the two analyses, namely that of total looking time. Although the latter measure is the one that is used more often, the former one—first look duration—could be more reliable because it discards any potentially "accidental" looking behavior that is not necessarily triggered by the auditory stimulation at hand (see Lewis et al., 1966). Having detected age-specific preferences for native versus non-native rhythm in only one of two measures, it should be noted that even that significant interaction effect was not detected with sufficient power: the power of ~0.60 that we observed was below the desired 0.80, although it was comparable to previous infant preference studies (Bergmann et al., 2018). For the above reasons, a replication of the present experiment is needed. Such a follow-up experiment could use identical stimuli but perhaps a paradigm fully contingent on infant looking behavior, which could help resolve whether the effects detected here for first look duration are real.

The present findings underline the role of future cross-linguistic studies focusing on early discrimination and preferences particularly in languages that are not straightforwardly identifiable as belonging to one of the traditional rhythmic classes. Such languages, more numerous than has been acknowledged in developmental language research, are currently understudied. We believe that adding findings about rhythmically unclassified languages is essential for completing the picture on early language discrimination. It may show that the development of language discrimination in infants is best explained by perceptual salience: the larger the rhythmic contrast between languages (defined not in terms of a membership to putative rhythm classes but in terms of an acoustic-perceptual distance), the better and earlier the discrimination.

ACKNOWLEDGEMENTS

This work was funded by grant 18-01799S from the Czech Science Foundation. NP and KC were also supported by Charles University grant Primus/17/HUM/19. We are grateful to Radka Klimičková, Kristýna Hrdličková and Veronika Ungrová for assistance with participant recruitment, testing, and data annotation, to Daniel Kopecký for help with stimulus preparation, and to Zuzana Oceláková for help with data organization and processing.

CONFLICT OF INTEREST

The authors declare no conflicts of interest with regard to the funding source for this study.

MATERIALS AND DATA AVAILABILITY

The stimuli, data, and analysis script are available at https://osf.io/zaj4y

ORCID

Nikola Paillereau https://orcid.org/0000-0002-3162-7231
Václav Jonáš Podlipský https://orcid.org/0000-0002-2173-4375
Filip Smolík https://orcid.org/0000-0003-4160-6124
Šárka Šimáčková https://orcid.org/0000-0002-9462-4169
Kateřina Chládková https://orcid.org/0000-0003-2009-1897

REFERENCES

Abboub, N., Nazzi, T., & Gervain, J. (2016). Prosodic grouping at birth. *Brain and Language*, *162*, 46–59. https://doi.org/10.1016/j.bandl.2016.08.002

Aldridge, M. A., Braga, E. S., Walton, G. E., & Bower, T. G. R. (1999). The intermodal representation of speech in newborns. *Developmental Science*, *2*(1), 42–46. https://doi.org/10.1111/1467-7687.00052

Arvaniti, A. (2012). The usefulness of metrics in the quantification of speech rhythm. *Journal of Phonetics*, *40*(3), 351–373. https://doi.org/10.1016/j.wocn.2012.02.003

Bates, D., Mächler, M., Bolker, B., & Walker, S. (2015). Fitting linear mixed-effects models using lme4. *Journal of Statistical Software*, *67*(1), 1–48. https://doi.org/10.18637/jss.v067.i01

Bergmann, C., & Cristia, A. (2016). Development of infants' segmentation of words from native speech: A meta-analytic approach. *Developmental Science*, *19*(6), 901–917. https://doi.org/10.1111/desc.12341

Bergmann, C., Rabagliati, H., & Tsuji, S. (2019). *What's in a looking time preference? Understanding statistical sign errors in infant cognition research*, https://doi.org/10.31234/osf.io/6u453

Bergmann, C., Tsuji, S., Piccinini, P. E., Lewis, M. L., Braginsky, M., Frank, M. C., & Cristia, A. (2018). Promoting replicability in developmental research through meta-analyses: Insights from language acquisition research. *Child Development*, *89*(6), 1996–2009. https://doi.org/10.1111/cdev.13079

Boersma, P., & Weenink, D. (2020). *Praat: doing phonetics by computer [Computer program]*. Version 6.1.09, Retrieved from http://www.praat.org/. Accessed 26 January 2020.

Bosch, L. (1998). Bilingual exposure and some conse- quences on native language recognition processes at four months. In P. K. Kuhl, & L. Crum (Eds.), *Proceedings of the Joint Meeting of the Acoustical Society of America and the International Conference on Acoustics* (pp. 1599–1600). Acoustical Society of America.

Bosch, L., & Sebastián-Gallés, N. (1997). Native-language recognition abilities in 4-month-old infants from monolingual and bilingual environments. *Cognition*, *65*(1), 33–69. https://doi.org/10.1016/s0010-0277(97)00040-1

Bosch, L., & Sebastián-Gallés, N. (2001). Evidence of early language discrimination abilities in infants from bilingual environments. *Infancy*, *2*(1), 29–49. https://doi.org/10.1207/s15327078in0201_3

Butler, J., Floccia, C., Goslin, J., & Panneton, R. (2011). Infants' discrimination of familiar and unfamiliar accents in speech. *Infancy*, *16*(4), 392–417. https://doi.org/10.1111/j.1532-7078.2010.00050.x

Byers-Heinlein, K., Burns, T. C., & Werker, J. F. (2010). The roots of bilingualism in newborns. *Psychological science*, *21*(3), 343–348. https://doi.org/10.1177/0956797609360758.

Cristia, A., Seidl, A., Vaughn, C., Schmale, R., Bradlow, A., & Floccia, C. (2012). Linguistic processing of accented speech across the lifespan. *Frontiers in Psychology*, *3*, 479. https://doi.org/10.3389/fpsyg.2012.00479

Dankovičová, J., & Dellwo, V. (2007). Czech speech rhythm and the rhythm class hypothesis. In *Proc. 16th Int. Congr. Phonetic Sciences, 6–10 August 2007* (pp. 471–474). Universität des Saarlandes.

Dellwo, V. (2006). Rhythm and speech rate: A variation coefficient for deltaC. In: P. Karnowski, & I. Szigeti (Eds.), *Language and Language-Processing: Proceedings of the 38th Linguistics Colloquium* (pp. 231–241). Peter Lang.

Diehl, M., Panneton, R., Kitamura, C., & Burnham, D. (2006). *Six-month-olds' perception of native speech accent*. Paper presented at the annual meeting of the XVth Biennial International Conference on Infant Studies, Westin Miyako, Kyoto, Japan.

Gasparini, L., Langus, A., Tsuji, S., & Boll-Avetisyan, N. (2020). *Quantifying the role of rhythm in infants' language discrimination abilities: A meta-analysis*. Preprint, https://doi.org/10.31219/osf.io/rmn5x

Gelman, A., & Carlin, J. (2014). Beyond power calculations. *Perspectives on Psychological Science*, *9*(6), 641–651. https://doi.org/10.1177/1745691614551642

Gervain, J. (2018). Gateway to language: The perception of prosody at birth. In H. Bartos, M. den Dikken, Z. Bánréti, & T. Váradi (Eds.), *Boundaries crossed, at the interfaces of morphosyntax, phonology, pragmatics and semantics* (pp. 373–384). Springer. https://doi.org/10.1007/978-3-319-90710-9_23

Gleitman, L., & Wanner, E. (1982). Language acquisition: The state of the state of the art. In E. Wanner, & L. Gleitman (Eds.), *Language acquisition: The state of the art* (pp. 3–48). Cambridge University Press.

Grabe, E., & Low, E. L. (2002). Durational variability in speech and the rhythm class hypothesis. In C. Gussenhoven, & N. Warner (Eds.), *Papers in laboratory phonology* (pp. 515–546). Mouton de Gruyter.

Green, P., & MacLeod, C. J. (2016). SIMR: An R package for power analysis of generalized linear mixed models by simulation. *Methods in Ecology and Evolution*, *7*, 493–498. https://doi.org/10.1111/2041-210X.12504

Hirsh-Pasek, K., Kemler Nelson, D., Jusczyk, P. W., Wright, K., Druss, B., & Kennedy, L. J. (1987). Clauses are perceptual units for young infants. *Cognitive Psychology*, *24*, 252–293.

Hood, B. M., & Atkinson, J. (1993). Disengaging visual attention in the infant and adult. *Infant Behavior and Development*, *16*(4), 405–422. https://doi.org/10.1016/0163-6383(93)80001-o

Houston-Price, C., & Nakai, S. (2004). Distinguishing novelty and familiarity effects in infant preference procedures. *Infant and Child Development*, *13*(4), 341–348. https://doi.org/10.1002/icd.364

Jusczyk, P. W., Friederici, A. D., Wessels, J. M. I., Svenkerud, V. Y., & Jusczyk, A. M. (1993). Infants′ sensitivity to the sound patterns of native language words. *Journal of Memory and Language*, *32*(3), 402–420. https://doi.org/10.1006/jmla.1993.1022

Jusczyk, P. W., Hirsh-Pasek, K., Kemler Nelson, D., Kennedy, L., Woodward, A., & Piwoz, J. (1992). Perception of acoustic correlates of major phrasal units by young infants. *Cognitive Psychology*, *24*, 252–293.

Jusczyk, P. W., & Thompson, E. (1978). Perception of a phonetic contrast in multisyllabic utterances by 2-month-old infants. *Perception and Psychophysics*, *23*(2), 105–109. https://doi.org/10.3758/bf03208289

Keij, B. (2017). *Rhythm & cues: Rhythmic structure and segmentation in early language acquisition* (Doctoral dissertation). LOT. ISBN 978-94-6093-233-5.

Kipp, M. (2014). ANVIL: A universal video research tool. In J. Durand, U. Gut, & G. Kristofferson (Eds.), *Handbook of corpus phonology* (pp. 420–436). Oxford University Press. Software version.

Kitamura, C., Panneton, R., & Best, C. T. (2013). The Development of language constancy: Attention to native versus nonnative accents. *Child Development*, *84*(5), 1686–1700. https://doi.org/10.1111/cdev.12068

Kitamura, C., Panneton, R., Deihl, M., & Notley, A. (2006). Attuning to the native dialect: when more means less. In P. Warren, & C. I. Watson (Eds.), *Proceedings of the 11th Australasian International Conference on Speech Science & Technology* (pp. 124–129). Australian Speech Science & Technology Association.

Kominsky, J. F. (2019). PyHab: Open-source real time infant gaze coding and stimulus presentation software. *Infant Behavior and Development*, *54*, 114–119. https://doi.org/10.1016/j.infbeh.2018.11.006

Kuznetsova, A., Brockhoff, P. B., & Christensen, R. H. B. (2017). lmerTest package: Tests in linear mixed effects models. *Journal of Statistical Software*, *82*(13), https://doi.org/10.18637/jss.v082.i13

Lewis, M., Kagan, J., & Kalafat, J. (1966). Patterns of fixation in the young infant. *Child Development*, *37*(2), 331–341. https://doi.org/10.2307/1126807

Lüdecke, D. (2018). ggeffects: Tidy data frames of marginal effects from regression models. *Journal of Open Source Software*, *3*(26), 772. https://doi.org/10.21105/joss.00772

May, L., Gervain, J., Carreiras, M., & Werker, J. F. (2017). The specificity of the neural response to speech at birth. *Developmental Science*, *21*(3), e12564. https://doi.org/10.1111/desc.12564

Mehler, J., & Christophe, A. (1995). Maturation and learning of language in the first year of life. In M. S. Gazzaniga (Ed.), *The cognitive neurosciences* (pp. 943–954). Bradford Books.

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. https://doi.org/10.1016/0010-0277(88)90035-2

Molnar, M., Gervain, J., & Carreiras, M. (2014). Within-rhythm class native language discrimination abilities of basque-spanish monolingual and bilingual infants at 3.5 months of age. *Infancy*, *19*(3), 326–337. https://doi.org/10.1111/infa.12041.

Moon, C., Cooper, R. P., & Fifer, W. P. (1993). Two-day-olds prefer their native language. *Infant Behavior and Development*, *16*(4), 495–500. https://doi.org/10.1016/0163-6383(93)80007-u

Mueller, J. L., Männel, C., & Friederici, A. D. (2015). Biological preconditions for language development. In J. D. Wright (Ed.), *International Encyclopedia of the Social and Behavioral Sciences* (2nd edition), (pp. 650–655). Elsevier. https://doi.org/10.1016/B978-0-08-097086-8.23117-6

Nazzi, T., Bertoncini, J., & Mehler, J. (1998). Language discrimination by newborns: Toward an understanding of the role of rhythm. *Journal of Experimental Psychology: Human Perception and Performance*, *24*(3), 756–766. https://doi.org/10.1037/0096-1523.24.3.756

Nazzi, T., Iakimova, G., Bertoncini, J., Fredonie, S., & Alcantara, C. (2006). Early segmentation of fluent speech by infants acquiring French: Emerging evidence for crosslinguistic differences. *Journal of Memory and Language*, *54*(3), 283–299. https://doi.org/10.1016/j.jml.2005.10.004

Nazzi, T., Jusczyk, P. W., & Johnson, E. K. (2000). Language discrimination by english-learning 5-month-olds: Effects of rhythm and familiarity. *Journal of Memory and Language*, *43*(1), 1–19. https://doi.org/10.1006/jmla.2000.2698

Peirce, J., Gray, J. R., Simpson, S., MacAskill, M., Höchenberger, R., Sogo, H., Kastman, E., & Lindeløv, J. K. (2019). PsychoPy2: experiments in behavior made easy. *Behavior Research Methods*, *51*(1), 195–203. https://doi.org/10.3758/s13428-018-01193-y

Phan, J., & Houston, D. M. (2009). Infant dialect discrimination. *The Journal of the Acoustical Society of America*, *125*(4), 2776. https://doi.org/10.1121/1.4784764

R Core Team (2019). *R: A language and environment for statistical computing*. https://www.R-project.org/

Rabagliati, H., Ferguson, B., & Lew-Williams, C. (2019). The profile of abstract rule learning in infancy: Meta-analytic and experimental evidence. *Developmental Science*, *22*(1), e12704. https://doi.org/10.1111/desc.12704

Ramus, F. (2002). Language discrimination by newborns: Teasing apart phonotactic, rhythmic, and intonational cues. *Annual Review of Language Acquisition*, *2*, 85–115. https://doi.org/10.1075/arla.2.05ram

Ramus, F., Hauser, M. D., Miller, C., Morris, D., & Mehler, J. (2000). Language discrimination by human newborns and by cotton-top tamarin monkeys. *Science*, *288*(5464), 349–351. https://doi.org/10.1126/science.288.5464.349

Ramus, F., Nespor, M., & Mehler, J. (1999). Correlates of linguistic rhythm in the speech signal. *Cognition*, *73*(3), 265–292. https://doi.org/10.1016/s0010-0277(99)00058-x

Sansavini, A., Bertoncini, J., & Giovanelli, G. (1997). Newborns discriminate the rhythm of multisyllabic stressed words. *Developmental Psychology*, *33*(1), 3–11. https://doi.org/10.1037/0012-1649.33.1.3

Sundara, M., Ngon, C., Skoruppa, K., Feldman, N. H., Onario, G. M., Morgan, J. L., & Peperkamp, S. (2018). Young infants' discrimination of subtle phonetic contrasts. *Cognition*, *178*, 57–66. https://doi.org/10.1016/j.cognition.2018.05.009

Thorson, J. (2018). The role of prosody in early word learning: Behavioral evidence. In P. Prieto, & N. Esteve-Gibert (Eds.), *The development of prosody in first language acquisition* (pp. 59–78). John Benjamins. https://doi.org/10.1075/tilar.23.04tho

Werker, J. F., & Gervain, J. (2013). Speech perception in infancy. *The Oxford Handbook of Developmental Psychology*, *1*, 908–925. https://doi.org/10.1093/oxfordhb/9780199958450.013.0031

Werker, J. F., & Tees, R. C. (2005). Speech perception as a window for understanding plasticity and commitment in language systems of the brain. *Developmental Psychobiology*, *46*(3), 233–251. https://doi.org/10.1002/dev.20060.

White, L., & Mattys, S. L. (2007). Calibrating rhythm: First language and second language studies. *Journal of Phonetics*, *35*(4), 501–522. https://doi.org/10.1016/j.wocn.2007.02.003

White, L., & Mattys, S. L. (2007a). Rhythmic typology and variation in first and second languages. In P. Prieto, J. Mascaró, & M.-J. Solé (Eds.), *Segmental and prosodic issues in romance phonology. Current issues in linguistic theory series* (pp. 237–257). John Benjamins.

White, L., Mattys, S. L., & Wiget, L. (2012). Language categorization by adults is based on sensitivity to durational cues, not rhythm class. *Journal of Memory and Language*, *66*(4), 665–679. https://doi.org/10.1016/j.jml.2011.12.010

White, L., Payne, E., & Mattys, S. L. (2009). Rhythmic and prosodic contrast in Venetan and Sicilian Italian. In M. Vigario, S. Frota, & M. J. Freitas (Eds.), *Phonetics and phonology: Interactions and interrelations* (pp. 137–158). John Benjamins.

How to cite this article: Paillereau N, Podlipský VJ, Smolík F, Šimáčková Š, Chládková K. The development of infants' sensitivity to native versus non-native rhythm. *Infancy*. 2021;26:422–440. https://doi.org/10.1111/infa.12395