Development of Auditory and Spontaneous Movement Responses to Music over the First Postnatal Year

Trinh Nguyen[1,2,3*], Félix Bigand[1], Susanne Reisner[2,4], Atesh Koul[1], Roberta Bianco[1,5], Gabriela Markova[6], Stefanie Hoehl[2], & Giacomo Novembre[1*]

[1]Neuroscience of Perception and Action Lab, Italian Institute of Technology, Rome, Italy
[2]Department of Developmental and Educational Psychology, University of Vienna, Vienna, Austria
[3]Department of Developmental and Biological Psychology, Heidelberg University, Heidelberg, Germany
[4]Doctoral School Cognition, Behavior and Neuroscience, University of Vienna, Austria
[5]Department of Translational Research & New Technologies in Medicine and Surgery, University of Pisa, Pisa, Italy
[6]Institute for Early Life Care, Paracelsus Medical University, Salzburg, Austria.

*Corresponding Authors. Trinh Nguyen, Viale Regina Elena 291, 00161 Rome. Trinh.nguyen@iit.it. Giacomo Novembre, Viale Regina Elena 291, 00161 Rome. Giacomo.novembre@iit.it

Acknowledgements

We are grateful to Eluisa Nimpf, Flavia Arnese, Larissa Reitinger, Josefine Schürholz and Liesbeth Forsthuber for their support in data acquisition, processing, and video coding. Additionally, we thank all families who participated in the study and the Department of Obstetrics and Gynaecology of the Vienna General Hospital for supporting our participant recruitment.

Funding

This research has received funding from the European Research Council awarded to Giacomo Novembre (MUSICOM, 948186), the Austrian Science Fund (FWF) DK Grant "Cognition & Communication 2" W1262-B29 awarded to Stefanie Hoehl (10.55776/W1262) and has been supported by the Brain and Machines Flagship Programme of the Italian Institute of Technology. Trinh Nguyen and Roberta Bianco also acknowledge the support of Horizon Europe's Marie Skłodowska-Curie Actions (SYNCON, 101105726; PHYLOMUSIC, 101064334).

Credits

Trinh Nguyen: Conceptualization, Methodology, Software, Formal Analysis, Investigation, Data curation, Visualization, Writing - Original draft, Writing – Review & Editing, Project administration; **Félix Bigand**: Software, Validation, Data curation, Formal Analysis, Writing – Review & Editing. **Susanne Reisner**: Investigation, Data curation, Writing – Review & Editing, Project administration. **Atesh Koul**: Software, Writing – Review & Editing; **Roberta Bianco**: Validation, Software, Writing - Original draft, Writing – Review & Editing; **Gabriela Markova**: Conceptualization, Data curation, Writing - Review & Editing; **Stefanie Hoehl**: Conceptualization, Methodology, Resources, Writing - Review & Editing, Supervision, Funding acquisition; **Giacomo Novembre**: Conceptualization, Methodology, Resources, Writing - Original draft, Writing - Review & Editing, Supervision, Funding acquisition.

Data availability

The data and stimuli reported in this manuscript are available in the following repository: https://doi.org/10.48557/DCSCFO. The codes used for analyses and figures are available on Github: https://github.com/tnguyen1992/tinydancer.

Abstract

Humans across cultures not only share the ability to recognise music but also respond to it through movement. While the sensory encoding of music is well-studied, when and how infants naturally start moving to music is largely unexplored. This study simultaneously investigates infants' neural (auditory) responses and spontaneous movements to music during the first postnatal year. Neural activity (EEG) and body kinematics (markerless pose estimation) were recorded from 79 infants (aged 3, 6, and 12 months) listening to refrains of children's music, along with shuffled, high-pitched, and low-pitched versions of the same songs. Neural data revealed that, across all ages, infants exhibit enhanced auditory responses to music compared to shuffled music, indicating that auditory encoding of music emerges early in development. Movement data revealed a different outcome. While coarse auditory-motor coupling is present at all ages, more complex structured movement patterns emerge in response to music only by 12 months. Notably, no age group demonstrated evidence of coordinated movements to music. Additionally, enhanced auditory responses to high vs low pitch were only evident at 6 months, while infants' movements were better predicted by high-pitched compared to low-pitched music at all ages. This study provides initial insights into how the developing brain gradually transforms music into spontaneous movements of increasing complexity.

Keywords: Music, EEG, movement, development, infancy, pitch

Introduction

Musicality – the biological predisposition to perceive, appreciate, and produce music (Honing, 2018; Trehub, 2003) – is increasingly recognised as a fundamental aspect of human nature. Numerous accounts suggest that engaging with music through movement is at the core of musicality (Honing et al., 2015; Schachner et al., 2009; Trehub et al., 2015). Functionally, such engagement can be broken down into two fundamental components of neurocognitive development: the ability to perceive and recognise music (*sensory component*), and the ability to produce movement responses that are temporally aligned with the musical structure, from coordinated vocalizations and percussive actions up to complex dance moves (*motor component*; Brown, 2022; Trehub, 2003a; Trevarthen, 1999). Despite this inherent predisposition toward music, the developmental trajectory of infants' musicality remains largely unknown (see Nguyen et al., 2023, for a review). While there is increasing research on infant music perception, including controlled manipulations of select musical features, we know less about the translation of perception into action, namely the ontogenesis of infants' spontaneous movements to music (see Fujii et al., 2014; Nguyen, Reisner, et al., 2023; Zentner & Eerola, 2010). Furthermore, making our understanding of music-driven motor engagement even more incomplete, no studies to date have looked at both brain activity and spontaneous body movements simultaneously, especially during the first year of life. Accordingly, how the processing of music and its features is transformed into organised motor responses remains underexplored.

The *sensory component* of musicality, namely music perception, can be measured using electroencephalography (EEG), specifically by recording cortical auditory evoked potentials (event-related potentials [ERP]). One of these responses is the infantile P1, a phase-locked EEG positivity peaking around 200-300 ms after an auditory stimulus (Chen et al., 2016; Kushnerenko et al., 2002; Wunderlich et al., 2006). The infantile P1 has been observed in response to both musical notes and speech segments. Auditory evoked potentials, when elicited isochronously, can also be captured using frequency domain analyses (Damsma et al., 2024; Novembre & Iannetti, 2018) such as auditory steady-state responses (ASSR), which are also called steady-state evoked potentials (SSEP, e.g., Cirelli et al., 2016; Nave et al., 2022). These neural responses can provide insight into the developing auditory system and its ability to encode musical structure. Using these neurophysiological measures, prior research has shown that newborns and infants are sensitive to beat structure, pitch deviants and tone interval regularities (Bianco et al., 2025; Edalati et al., 2023; Háden et al., 2009, 2015, 2022; Stefanics et al., 2009; Winkler et al., 2009). Despite these promising results, the

neurophysiology of early music processing – particularly its developmental trajectory – remains not fully understood. Here, our primary goal is to investigate infants' neural encoding of music utilizing both ERP and ASSR approaches to characterize how such neural responses change across the first year of life.

Another component of musicality is the capacity to move to music (*motor component;* Brown, 2022; Fitch, 2015; Honing et al., 2015; Trehub et al., 2015). This capacity is linked to infants not only recognising musical structure but also moving their bodies in response to it. Even though this capacity appears to develop precociously, as evidenced by the fact that even 28-35-week-old foetuses move to music (Kisilevsky et al., 2004), very few studies have systematically examined music-driven spontaneous body movements in infants. An influential paper by Zentner and Eerola (2010) reported that infants across a large age range (from 5 to 24 months) showed more spontaneous rhythmic movements in response to classical music and children's music compared to infant-directed speech. Importantly, their movements were not synchronized with the musical input, even though a small degree of tempo flexibility was observed (i.e., faster musical tempi evoked relatively faster movement periodicities). The lack of synchrony between music and body movements has also been reported in younger (i.e., 3-4 months old) infants listening to popular music (Fujii et al., 2014). Further, another study testing 7-month-old infants reported more movement in response to (sung) playsongs compared to lullabies but did not assess movement synchrony (Nguyen, Reisner, et al., 2023). Despite these initial investigations, it remains unclear when infants begin to move in response to music, which specific movements are evoked, and when these movements become coordinated with the music. Moreover, a critical limitation in existing research is the lack of a control condition to determine whether these movements are driven specifically by musical structure or reflect general motor activity in response to auditory input. As a second goal, this study is the first to systematically test the gradual development of music-induced movements in different age groups across the first year of life.

Music engages both sensory and motor systems, yet different musical features may differentially shape infants' engagement with music. While rhythm has been widely studied in early music cognition, pitch is another salient acoustic cue that could play a role in auditory-motor engagement, particularly in infancy. High pitch is a defining feature of infant-directed speech (Fernald & Simon, 1984), among other features such as exaggerated intonation, slower tempo, and simplified vocabulary (Fernald & Kuhl, 1987; Kuhl & Meltzoff, 1982). Similarly, infants most frequently listen to music characterized by high pitch (Costa-Giomi & Sun, 2016; Nakata & Trehub, 2011). Reflecting its prominence, high pitch is

found to be one of the most prominent features thought to effectively capture (Conrad et al., 2011; Eckerdal & Merker, 2009; Trainor, 1996; Trainor & Zacharias, 1998) and guide infants' attention (Lense et al., 2022; Trainor & Desjardins, 2002). On the neural level, infants are also better at encoding pitch deviances in the high voice of polyphonic music, thus showing *high voice superiority* from 3 months of age (Marie & Trainor, 2013, 2014). Taken together, these findings indicate that higher-pitch music would amplify infants' neural responses (i.e., sensory component) in comparison to lower-pitch music. On the other hand, we know that adults move more to music with greater energy in lower frequencies (Cameron et al., 2022; Stupacher et al., 2013, 2016; Van Dyck et al., 2013). Yet, it remains unknown whether low-pitch music elicits increased movement in infants, as it does in adults, or whether infants' attraction to high pitch also extends to enhance their motor responses. As a third goal, we thus investigate how musical pitch affects infants' sensory and motor components.

We presented infants, aged 3, 6, and 12 months, with instrumental refrains of children's songs (music), shuffled versions of the same songs (shuffled music), and transpositions of the songs that would either emphasize the melody (high pitch) or the bassline (low pitch). We recorded infants' neural activity using EEG and specifically extracted ERPs and ASSR as indices of infants' neural response to the various auditory stimuli. We also analysed spontaneous (full-body) movement kinematics using automated video-based motion tracking (DeepLabCut) and extracted *principal movements* using principal component analysis (see Fig. 1, c.f. Bigand et al., 2024, Toiviainen et al., 2010). By adopting a cross-sectional design, we aimed to characterize the maturation of both auditory and movement responses across infancy. We hypothesized that auditory responses would be enhanced when triggered by music compared to shuffled music. This hypothesis was based on the notion that musical structure, notably eroded in the shuffled musical stimuli, is essential to attract infants' attention towards predictable events (Kouider et al., 2015; Lense et al., 2022). Similarly, based on previous evidence comparing movement responses to music vs speech and silence (Fujii et al., 2014; Zentner & Eerola, 2010), we expected the presence of musical structure to increase the likelihood of spontaneous movements in response to music compared to shuffled music, but we did not have a specific hypothesis about which particular movements would be produced. We further hypothesized that infants would show enhanced neural responses to high- compared to low-pitch music and explored co-occurring differences in spontaneous movements. Generally, we aimed to characterize the maturation of both auditory and motor responses as infants get older. By studying both sensory and motor components of musicality, we aimed to deepen our understanding of when and how infants

learn to transform what they perceive into spontaneous movements, eventually leading to the emergence of synchronization to music (Brown, 2022; Fitch, 2015; Honing et al., 2015; Patel & Iversen, 2014).

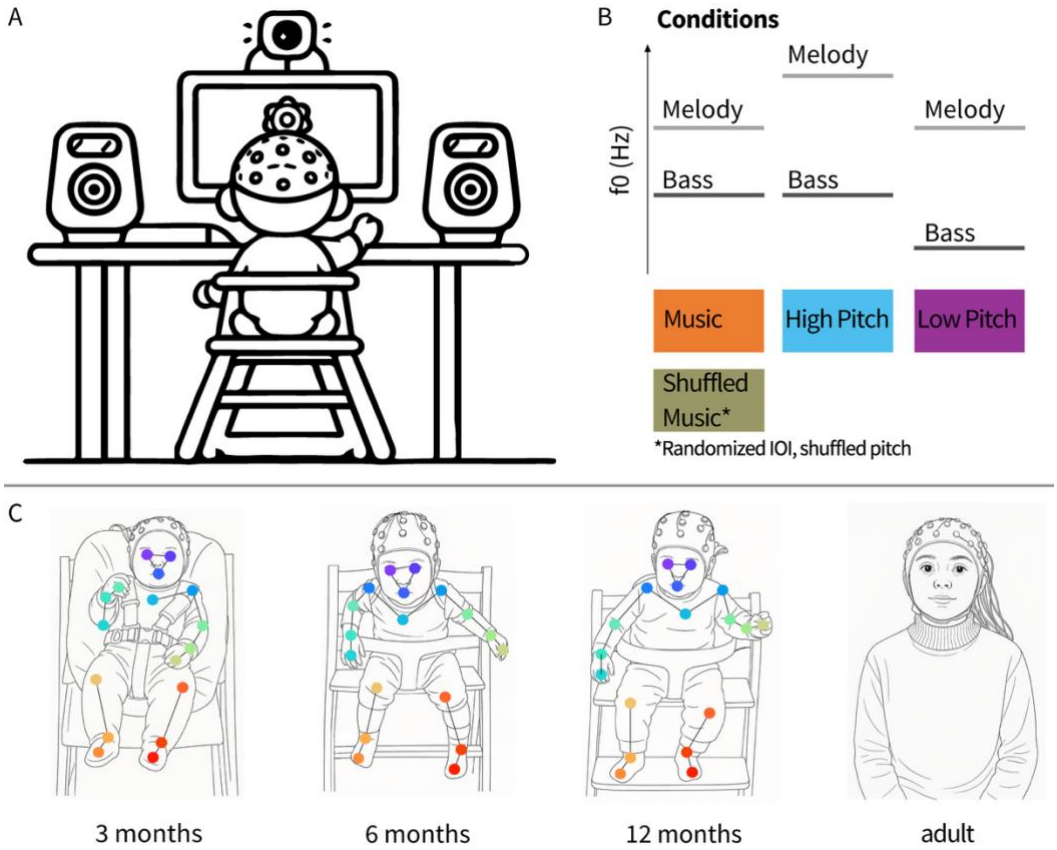

Figure 1. *Overview of the procedure (A), experimental conditions (B), and participant sample (C). (A) Infants sat in front of a screen with speakers on each side. The screen showed slowly blossoming flowers to attract infants' attention. Caregivers (not shown) sat behind the infants and wore noise-cancelling headphones. (B) Infants listened to polyphonic auditory stimuli consisting of a melody and a bassline in four different conditions. The music condition included two children's songs. The shuffled music condition included versions of the songs used in the music condition that were shuffled in pitch and randomized in inter-onset intervals (IOI). Stimuli belonging to the music and shuffled music conditions had the same pitches. In the high-pitch condition, the melody was shifted one octave higher than in the music condition. In the low-pitch condition, the bassline was shifted one octave lower than in the music condition. Hence, the two voices composing the high-pitch condition were one octave higher than those composing the low-pitch condition. (C) The sample included infants at 3 months (N=26), 6 months (N=26), 12 months (N=27), and an adult control sample (N=26). The dots overlaying the images represent the body parts whose movements were tracked using video-based kinematic analysis.*

Results

This study included EEG and movement measurements taken from 79 infants in the first postnatal year, as well as EEG measurements taken from a control sample of 26 adults. Participants were exposed to two polyphonic children's songs featuring a melody and a bassline, and their manipulated versions (see Table 1 and S1 for an acoustic characterization of the musical stimuli). We investigated neural responses and movement responses to music vs shuffled music (a control condition in which we shuffled the melody and randomized the inter-onset intervals (IOI) of the music). Additionally, we contrasted neural and movement responses to high- vs low-pitch music. Music vs shuffle conditions (manipulation of structure but not pitch height) and high pitch vs low pitch conditions (manipulation of pitch height but not structure) are contrasted separately to avoid comparing conditions differing in more than one variable. Below, we begin by reporting the neural results, followed by the movement results.

Table 1. Acoustic characteristics of the musical stimuli

Song name (Condition)	***TEMPO M±SD (BPM)***	***IOI RANGE (MS)***	***IPI M±SD (SEMITONES)***	***IPI RANGE (SEMITONES)***	***BASS NOTE AND PITCH RANGE (HZ)***	***MELODY NOTE AND PITCH RANGE (HZ)***
HOPP (MUSIC)	0.422±0.067	0.222-0.444	0.040±4.585	-10 - 7	E2 to E3 (82.41 – 164.82 Hz)	A3 to A4 (220.00 – 440.00 Hz)
HOPP (SHUFFLED MUSIC)	0.422±0.187	0.112-0.666	0.100±4.782	-7 - 12	E2 to E3 (82.41 – 164.82 Hz)	A3 to A4 (220.00 – 440.00 Hz)
HOPP (HIGH PITCH)	0.422±0.067	0.222-0.444	0.040±4.585	-10 - 7	E2 to E3 (82.41 – 164.81 Hz)	A4 to A5 (440.00 – 880.00 Hz)
HOPP (LOW PITCH)	0.422±0.067	0.222-0.444	0.040±4.585	-10 - 7	E1 to E2 (41.20 – 82.41 Hz)	A3 to A4 (220.00 – 440.00 Hz)
LOLA (MUSIC)	0.444±0.000	0.444-0.444	0.149±2.836	-7 - 7	E2 to E3 (82.41 – 164.82 Hz)	A3 to A4 (220.00 – 440.00 Hz)
LOLA (SHUFFLED MUSIC)	0.444±0.182	0.222-0.665	0.234±4.864	-12 - 11	E2 to E3 (82.41 – 164.82 Hz)	A3 to A4 (220.00 – 440.00 Hz)
LOLA (HIGH PITCH)	0.444±0.000	0.444-0.444	0.149±2.836	-7 - 7	E2 to E3 (82.41 – 164.81 Hz)	A4 to A5 (440.00 – 880.00 Hz)
LOLA (LOW PITCH)	0.444±0.000	0.444-0.444	0.149±2.836	-7 - 7	E1 to E2 (41.20 – 82.41 Hz)	A3 to A4 (220.00 – 440.00 Hz)

Note. M = mean, SD = standard deviation, BPM = beats per minute, IOI = inter-onset interval, IPI = inter-peak interval, Hz = hertz (frequency). IOI and IPI were computed across melody and bass notes. Hopp refers to the Hungarian playsong ("Hopp Juliska"), Lola refers to the Spanish playsong ("La vaca lola").

EEG: Event-related potentials (ERP)

To examine the neural responses, we analysed the average ERPs time-locked to the acoustic tone onsets of the bassline notes of the auditory stimuli (*Fig. 2*). Adults' ERPs, which served as a benchmark to interpret infants' responses, included an early positivity peaking at 37 ms post-stimulus (so-called "P50", here reaching an average amplitude of 1.05 µV), followed by a later negativity peaking at 87 ms post-stimulus (so-called "N100", here reaching an amplitude of -0.43 µV) and a second positivity peaking at 158 ms post-stimulus (so-called "P200", here reaching an amplitude of 0.85 µV). This triphasic EEG pattern has been widely observed in adults in response to fast-rising auditory stimuli and across different contexts (Novembre et al., 2018; Pratt et al., 2008; Remijn et al., 2014; Somervail et al., 2021). Cluster-based permutation analyses, contrasting the ERPs elicited by music vs shuffled music, revealed that the amplitude of the P50 – hereafter referred to as P1 – was larger in response to music compared to shuffled music, particularly within the time range comprised between -17 and 58 ms post-stimulus (*cluster-t*=366.16, *p*=.016). Similarly, the amplitude of the following P200 - hereafter referred to as P2 – was larger in response to music than to shuffled music, particularly within the time range of 114 to 190 ms post-stimulus (*cluster-t*=395.42, *p*=.016). Both P1 and P2 responses were observed over fronto-central electrodes, showing a medial distribution, in line with previous literature (e.g., Lijffijt et al., 2009).

All infants' ERPs showed a P1 response, while a P2 response was observed only in 12-month-old infants, albeit with a lower amplitude than the P1. The P1 latency decreased ($X^2(2)=391.25$, $p<.001$), and its amplitude increased ($X^2(2)=8.59$, $p=.014$) with age (*Fig. 2*, left column). Importantly, and in line with the adults' data, all infant groups exhibited enhanced P1 amplitudes in response to music compared to shuffled music. Actually, across all groups, shuffled music did not elicit clear ERPs as the ones elicited by music. Cluster-based permutation (n_{Perm}=1000) testing revealed that 3-month-old infants' P1 amplitude was enhanced between 177 and 305 ms post-stimulus (*cluster-t*=1111.90, *p*=.002). Within this window, the P1 peaked at 212 ms and reached an average amplitude of 1.8 µV. The topography included a frontocentral cluster with a slight right lateralization. In 6-month-old infants, the amplitude of the P1 was enhanced between 116 and 284 ms post-stimulus

(*cluster-t*=1401.60, *p*=.002), peaking at 165 ms and reaching an amplitude of 2.8 µV. The topography included a few (centro-) parietal electrodes in addition to several frontocentral electrodes, with a bilateral activation. In 12-month-old infants, the amplitude of the neural response to music was enhanced in a two-peak cluster (*cluster-t*=1416.30, *p*=.002). The first peak, an infantile P1, occurred between 104 and 227 ms, peaked at 146 ms post-stimulus and reached an amplitude of 3.1 µV. Notably, 12-month-old infants exhibited an additional positivity, namely an infantile P2, possibly homologous to the P200 observed in adults. The P2 ranged from 307 to 325 ms post-stimulus and peaked at 316 ms, with an average amplitude of 1.026 µV. The topographies remained frontocentral but were more medial, similar to adults.

To rule out the possibility that increased neural responses in the music condition reflected differences in inter-tone temporal spacing (see IOI range in *Table 1*), we performed control analyses including only epochs from the shuffled music condition for which the prior or subsequent IOI exceeded the median IOI duration. The resulting ERPs in the shuffled music condition were highly similar to those reported above (compare *Fig. 2* with *Fig. S1*), and the difference between the music and shuffled music conditions remained significant across all age groups ($p<.05$). These findings indicate that the observed effects are not solely attributable to variations in inter-tone spacing but also reflect sensitivity to musical structure.

Next, we examined neural responses to the notes in the high- and low-pitch conditions (*Fig. 2,* right column). The morphology of both adults' and infants' ERPs was generally comparable to that elicited by the music condition. Cluster-based permutation (n_{Perm}=1000) testing revealed that the amplitude of the adults' ERPs was comparable across high- and low-pitch conditions ($p>.050$). This was also the case for both 3- and 12-month-old infants, but notably not for 6-month-old infants, who exhibited an enhanced P1 in response to high-pitch vs low-pitch conditions (*cluster-t*=763.84, *p*=.002). This enhanced positivity (178 and 332 ms) peaked at 204 ms and reached an average amplitude of 2.8 µV. Similar to the neural response elicited by the music condition, the topography included few (centro-) parietal electrodes in addition to several frontocentral electrodes, with a bilateral activation.

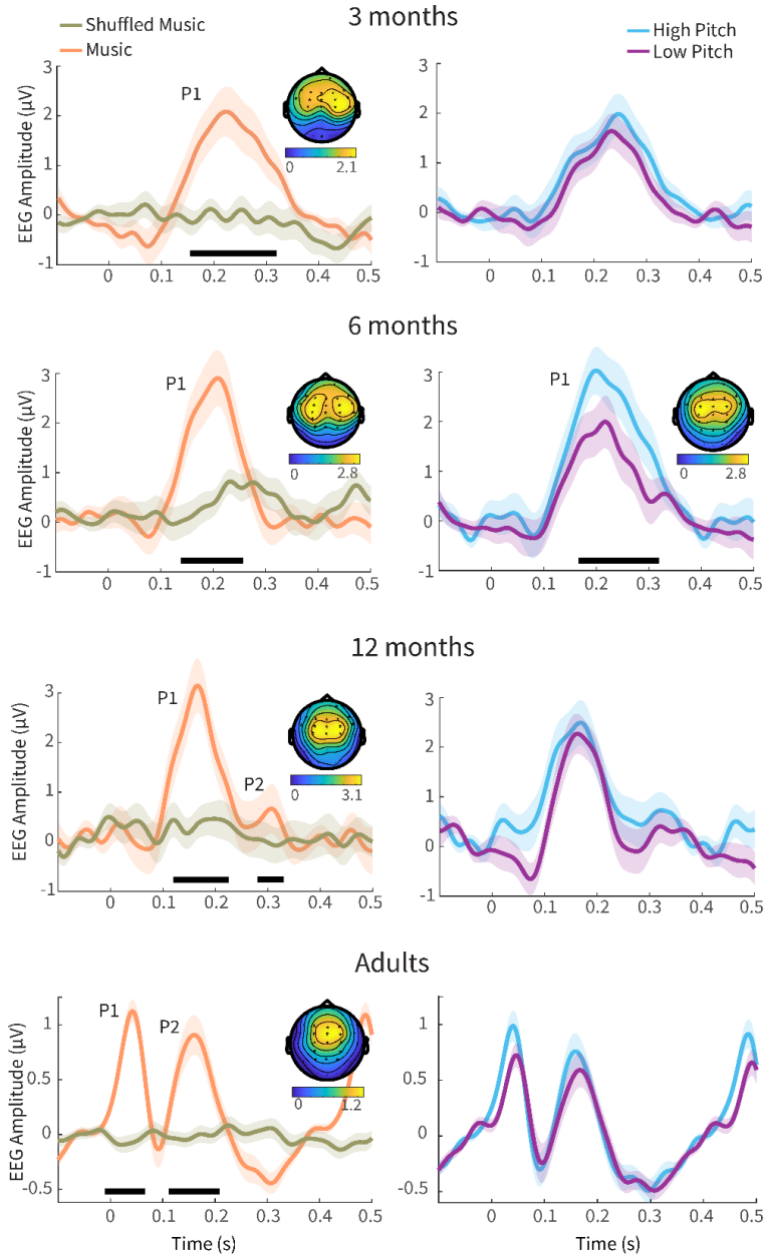

Figure 2. *Event-related potentials (ERPs) elicited by the notes comprised within the music (orange, left) vs shuffled music (khaki, left) as well as by the notes comprised within the high-pitch (light blue, right) vs low-pitch music (purple, right), across four groups of participants (plotted in ascending order of age, from top to bottom): 3-, 6-, 12-month-old infants (N=79) and adults (N=26). Grand-average ERPs are averaged across electrodes within the significant cluster of each age group in the music condition (except for pitch condition comparison in the 6-month-olds). Shaded areas indicate the standard error. ERPs show progressively shorter latencies with increasing age. All groups exhibited a P1 response, while only older infants (12-month-olds) and adults additionally exhibited a P2. Music elicited a larger P1 (and, when present, P2) amplitude compared to shuffled music, notably across all*

groups (time ranges associated with a significant difference are indicated by horizontal black lines). The topography of this neural response (averaged across the time window of the P1 cluster) in the music condition appears to shift more medially with increasing age. Colorbars beneath topography plots index EEG amplitude values.

EEG: Auditory Steady State Responses (ASSR)

Figure 3 shows bar plots indexing the relative power of ASSRs elicited by the auditory stimuli (power estimates were averaged across the electrodes comprised within the ERP clusters that were common to all age groups, i.e., FP2, F7, F3, Fz, F4, F8, FC7, FC3, FCz, FC4, FC8, C3, Cz, C4; see *Fig. 2*).

Our frequency of interest was 2.25 Hz, matching the musical beat as well as the presentation rate of the majority of the notes (15 out of 16 notes) in the auditory sequences (see methods, section "Stimuli"). We used linear mixed models, including power estimates as the dependent variable to contrast the different conditions and age groups (fixed and interaction effects). Participants were modelled as random intercepts. Power estimates were generally higher in response to music as opposed to shuffled music. Model outputs indicated that the power of the ASSR elicited by the music condition was significantly higher than that elicited by the shuffled music condition in 3-month-old infants ($F(1,50)=7.82$, $p=.007$), 12-month-old infants ($F(1,52)=12.03$, $p=.001$) and adults ($F(1,50)=13.49$, $p<.001$); while it only reached a trend for significance in 6-month-old infants ($F(1,50)=2.95$, $p=.092$). ASSR power estimates were not different across high-pitch and low-pitch conditions across all age groups ($p>.240$). These results indicated that most groups tracked the periodicity of the musical beat, leading to an enhancement of power in ASSR at a frequency matching the musical beat. This is generally in line with the ERP results showing stronger neural responses to music than shuffled music (even though 6-month-old infants only showed a marginal difference).

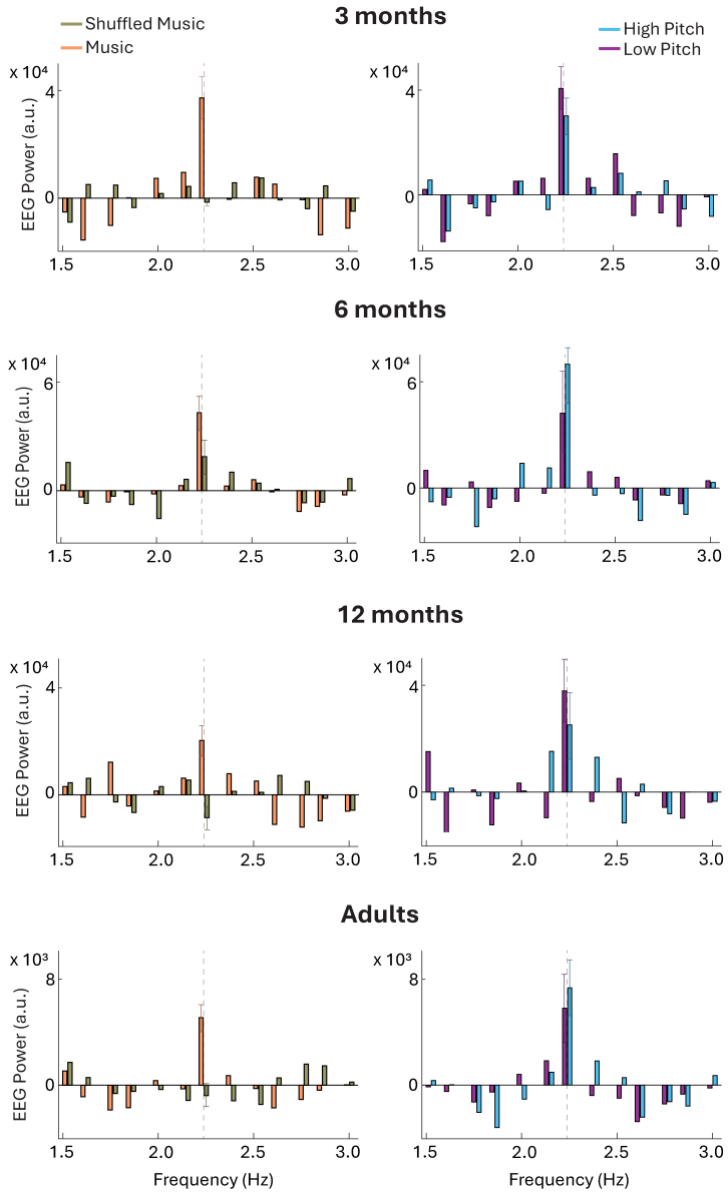

Figure 3. *Relative EEG Power (arbitrary unit [a.u.], y-axis) of the auditory steady-state responses (ASSR) elicited by music versus shuffled music (orange and khaki, left), and high-pitch versus low-pitch musical stimuli (blue and purple, right), across four groups of listeners: 3-month-olds (first row), 6-month-olds (second row), 12-month-olds (third row), and adults (fourth row). ASSR power estimates at the frequency (x-axis) matching the musical beat (2.25 Hz, highlighted by vertical dashed lines and including standard error bars) were statistically higher when elicited by music compared to shuffled music across nearly all participant groups (i.e., all but 6-month-olds). High- and low-pitch stimuli evoked similar ASSR (at 2.25 Hz). These results generally align with the ERP results (Fig. 2) across most infant groups and adults, except for 6-month-old infants for whom differences across conditions were either trending (music vs shuffled) or not significant (high vs low pitch).*

Extraction of Principal Movements and Estimation of Quantity of Movement

Using principal component analysis, we decomposed full-body kinematics into 10 Principal Movements (PMs), explaining 79.7 % of the whole kinematic variance. The PMs (depicted in Fig. 4) were reminiscent of common infant movements such as front-back rocking (PM1), side sway (PM2), proto-clapping (PM3), leg kicking (PM4), up-down rocking (PM5), arm pedalling (PM6), feet kicking (PM7), whole body wiggling (PM8), feet shuffling (PM9) and feet pedalling (PM10). Labels were assigned qualitatively, following visual inspection of the PMs.

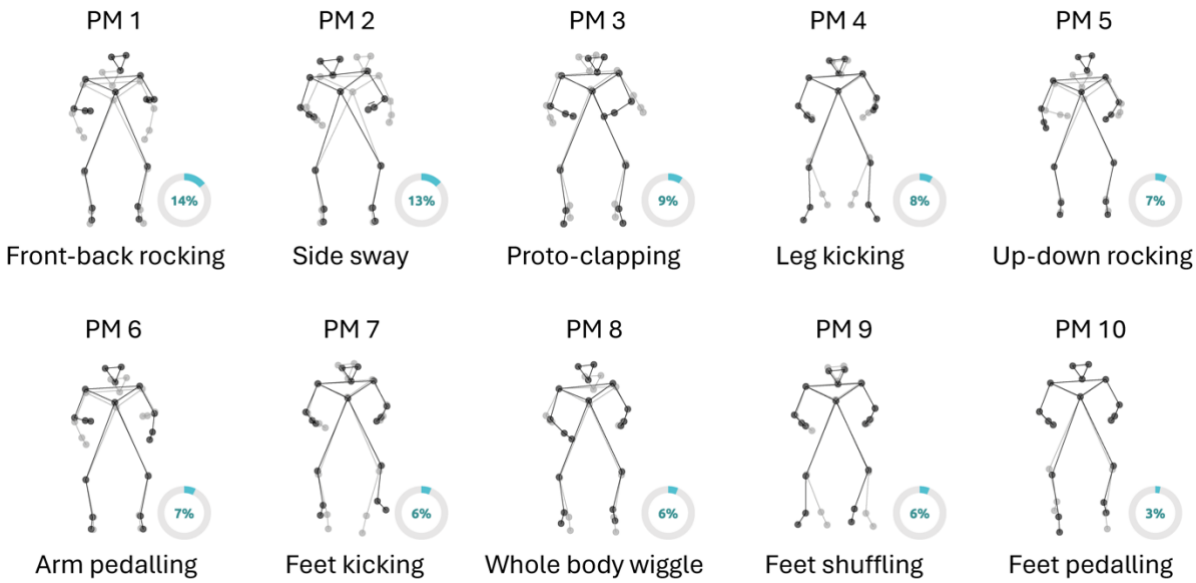

Figure 4. *Infants' principal movements (PMs). PMs are illustrated by showing the two most different body postures (min and max of the PM score, in grey and black, respectively) from the frontal perspective. The reader should interpret the PM as the kinematic displacement necessary to shift from one body posture (grey) to the other (black). Circle diagrams denote the proportion (%) of kinematic variance explained by each PM. Together, the ten PMs account for 79.7% of the total kinematic variance.*

Quantity of Movement (QoM) was estimated and compared across PMs, conditions, and groups (*Fig. 5*). QoM estimates were based on first-degree differentiation of the PM time series (see methods). A random intercept was included for each infant. Linear mixed effect modelling yielded a significant interaction between Condition and Age ($\chi^2(2)=16.76$, $p<.001$) indicating that only 12-month-old infants exhibited higher QoM in response to music compared to shuffled music across all PMs (i.e., when QoM was computed as an average across all PMs). Post-hoc contrasts comparing QoM averages between conditions revealed a

highly significant effect ($t(69.8)$=4.86, p<.001; see top left panel of Fig. 5 for QoM averaged across PMs).

Even though there was no interaction effect between Condition, Age and PMs, we still ran post-hoc comparisons to gain preliminary evidence about specific PMs driving the above-described effects. Results indicated that differences in 12-month-olds' QoM in response to music vs shuffled music were mostly driven by movements of the upper body and/or upper limbs. Specifically, front-back rocking (PM1), side sway (PM2), proto-clapping (PM3), up-down rocking (PM5) and arm pedalling (PM6) were linked with significantly higher QoM in response to music as opposed to shuffled music (PMs 1,2,3,5,6; *ps*<.050, corrected using the false-discovery rate). Contrarily, younger infants (3- and 6-month-olds) did not exhibit significantly different QoM in response to music vs shuffled music in any of the PMs (*ps*>.123). Further, when comparing infants' QoM to music at different pitches, we found no significant differences between conditions across PMs and age groups (*ps*>.295).

The model also yielded a significant interaction between PMs and Age ($\chi^2(18)$=181.575, p<.001), indicating that QoM generally increased with age but differently across PMs. Specifically, from 3 to 6 months, the PMs associated with a faster increase in QoM were front-back rocking (PM1), proto-clapping (PM3) and arm pedalling (PM6) ($t(95.4)$=2.34-2.60, p=.016-.032). From 3 to 12 months, front-back rocking (PM1), side sway (PM2), proto-clapping (PM3), up-down rocking (PM5) and arm pedalling (PM6) became more prevalent ($t(95.4)$=2.69-5.06, p=.001-.025). From 6 to 12 months, proto-clapping (PM3) became even more prevalent ($t(95.4)$=2.71, p=.012). Across the first year of life, infants seemed to consistently move their lower body while slowly increasing their capacity for upper-body and whole-body movements while seated.

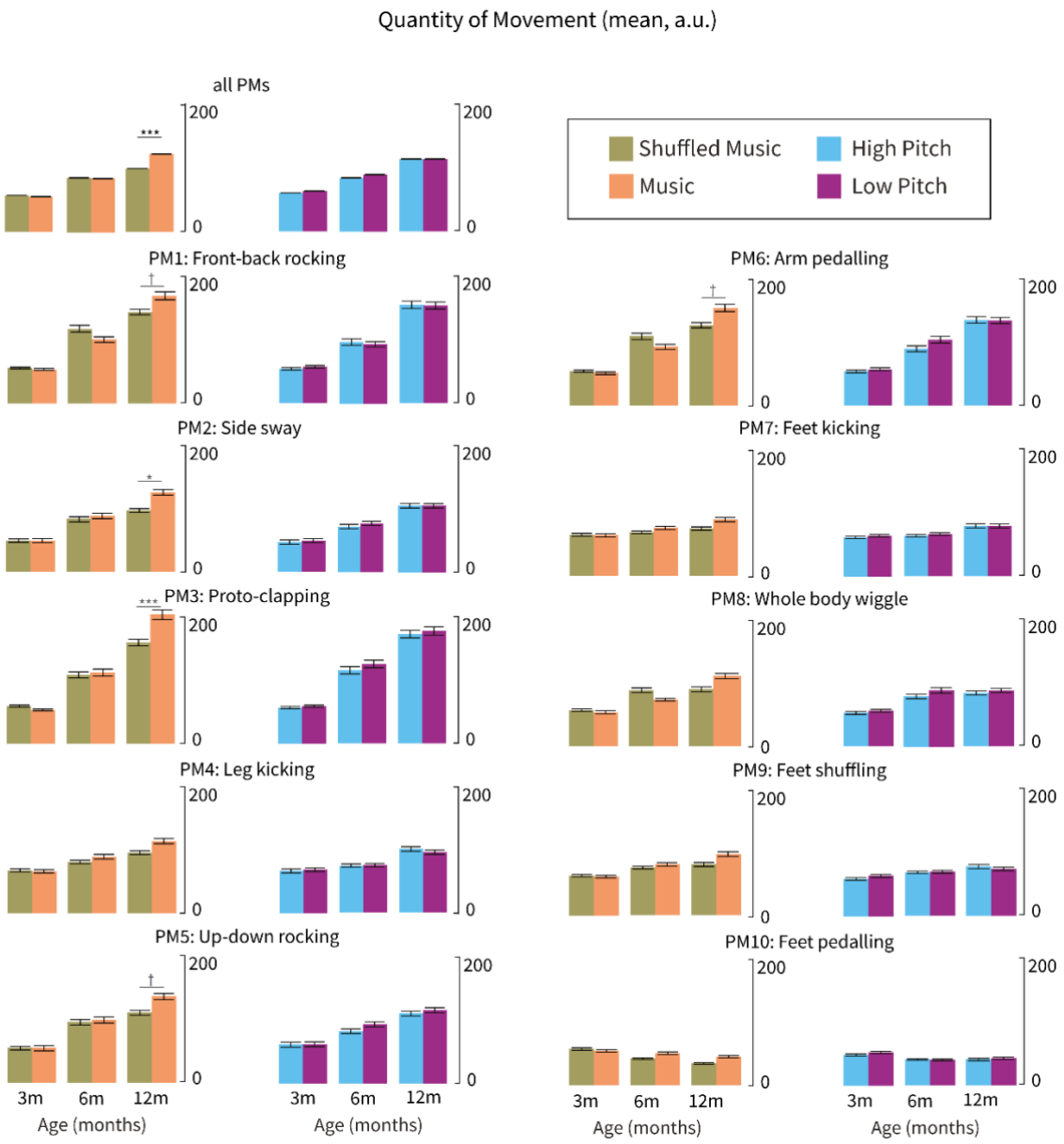

Figure 5. *Quantity of movement (mean, a.u.; [QoM]) elicited by music (orange) versus shuffled music (khaki) and high-pitch (blue) versus low-pitch music (purple) across different age groups (3-month-olds, 6-month-olds, 12-month-olds) and principal movements (PMs). Bar plots indicate the mean and standard error of QoM across different age groups, conditions, and PMs. Only twelve-month-old infants showed significantly increased QoM in response to music compared to shuffled music, specifically in PMs involving upper body movements (front-back rocking, side sway, proto-clapping, up-down rocking, and arm pedalling). No significant differences were observed between high- and low-pitch conditions. These results were also replicated in a supplementary analysis assessing differences in variance of (as opposed to mean) QoM (see Fig. S2 and Supplements for more details). † = p<.100, * = p<.050, ** = p<.010, *** = p<.001*

Movement: Granger Causality Analysis

Beyond looking at how much infants moved, we further investigated whether the spontaneous occurrence of infant movements could be explained by preceding changes in the intensity of the auditory stimuli. To do so, we used the sound envelope of the auditory stimuli (indexing changes in intensity over time) to predict infant movement velocity (time series representing changes in movement velocity, averaged across all PMs) and vice versa, using Granger-Causality analysis. Prediction estimates (Granger F-values) were yielded across different time lags, indexing the elapsed time necessary for a change in stimulus intensity to predict a change in movement velocity (i.e., highest F-values represent optimal prediction).

A preliminary analysis (sanity check) showed that musical stimuli predicted subsequent movement velocity better than vice versa across all age groups (*Fig. 6*, top-right corner; $F(1)=94.62$, $p<.001$). To identify the optimal temporal lags predicting sound-driven changes in movement, we conducted bootstrapped t-tests corrected for multiple comparisons across time points. Results (*Fig. 6*, left) showed that musical stimuli were better predictors of movement than shuffled stimuli at 3 months (lags of 160-200 ms, $t(40)=3.55$-3.98, $p<.001$), 6 months (lags of 120-240 ms, $t(37)=2.99$-4.49, $p<=.001$) and 12 months (lags of 160-240 ms, $t(31.6)=4.95$-5.03, $p<.001$). Conversely, there was a drop in prediction at later time lags in all groups (~320-360 ms; $t(43.6)=3.88$-8.25, $p<.001$). Together, these results suggest that infants' movement was related to intensity changes in the music, but not in the shuffled music.

Results associated with the high- and low-pitch conditions yielded similar Granger F-values to the music condition (*Fig. 6*, right). Notably, prediction estimates were generally higher for the high-pitch condition as compared to the low-pitch condition, indicating that high-pitch music was a better predictor of movement than low-pitch music. The difference between these two conditions was significant in one time-window in the 3-month-old infants (80-200 ms, $t(39.26)=2.66$-3.81, $p<.002$), while it encompassed two time windows in the 6-month-olds (120-200 ms, $t(45.64)=3.14$-4.21, $p<=.004$, and 320-360 ms, $t(45.7)=3.32$-3.82, $p<=.004$) and 12-month-olds (120-160ms, $t(40.4)=3.20$-3.73, $p<=.002$ and 520-600 ms, $t(39.9)=3.13$-3.41, $p<.004$) – perhaps suggesting that this effect grows stronger with age. Finally, Granger Causality statistics stratified by each PM and age group are detailed in the Supplements and *Fig. S2*.

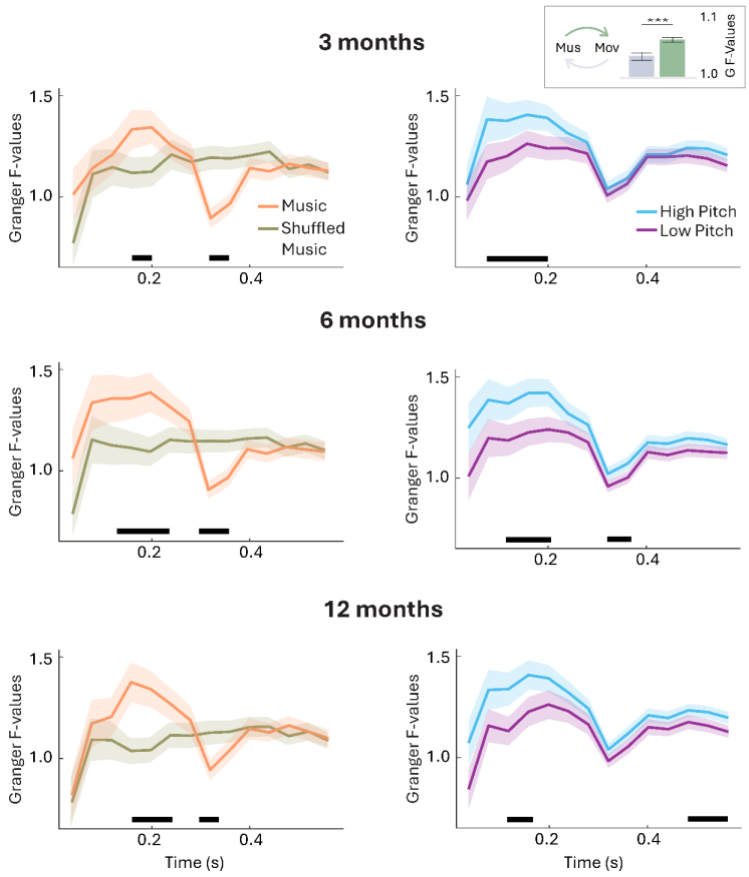

Figure 6. Music-driven movement *(Granger-Causality analysis). Top-right: A sanity check analysis showed that musical stimuli predicted subsequent movement velocity (green) better than vice versa (grey; p<.001). Left: Movement velocity was better predicted by music (orange) than by shuffled music (khaki), particularly with time lags of 160-200 ms (shaded areas indicate standard errors; horizontal black lines underline time ranges associated with a significant difference between conditions). Right: Movement was better predicted by high-pitch music (blue) compared to low-pitch music (purple).*

Movement: Phase-locked changes and periodicity

The Granger Causality analysis indicated that changes in music intensity drove movement in time, especially at a time lag comprised within a 200 ms delay. This result indirectly suggests that changes in music might evoke a phase-locked movement response. If so, we should be able to observe such phase-locked response when epoching movement data to peaks in the amplitude envelope of the auditory stimuli. To test this prediction, we ran supplementary event-related analyses on the movement velocity time series used for the previous analysis. Cluster-based permutation analyses revealed no significant clusters across age groups and conditions (*ps*>.050; see Supplements, *Fig. S3*), even though it should be

noted that 12-month-old infants exhibited a movement peak at ~200 ms, which was slightly but not significantly higher in response to music vs shuffled conditions. These results indicated that infants did not consistently exhibit phase-locked movement responses to musical events such as peaks in the amplitude envelope of the auditory stimuli.

Next, we also examined to what extent spontaneous movements were periodic and how so across conditions. This analysis builds upon the distribution of the highest coefficients yielded by auto-correlation analyses (similar to Zentner & Eerola, 2010). The results indicated that overall, spontaneous movements tended to be periodic, but such periodicity did not match the musical beat and was not different across conditions (see *Supplements*). In other words, movement periodicity did not result in coordination with music, and it was not modulated by whether infants were listening to music vs shuffled music or high- vs low-pitch music (*Fig. S4*). This result was further supported by an exploratory analysis in which we identified movement bursts by isolating epochs with high amplitude variation in the time series and excluding segments with low or no variation (c.f. Fujii et al., 2014). Together, these results suggest that while infants might generally exhibit rhythmic movements in response to sounds, the specificity of this behavioural response to music – and its coordination with musical structure – continues to develop beyond 12 months of age.

Discussion

This study examined the development of infants' neural and movement responses to music during the first year of life using a combination of neural measures, such as ERP and ASSR, alongside quantitative analyses of infants' movement. These approaches allowed us to explore both sensory and motor components of musicality, shedding light on how infants process and respond to musical stimuli. Below, we discuss 1) the development of neural auditory responses to music, 2) the emergence of music-driven movement patterns, and 3) the sensitivity of both components to pitch across early development.

Neural Responses to Music: Sensitivity to structured music across all ages

Much research suggests that the infants' auditory system is sensitive to structural features of music, such as timing and pitch regularities, from early on (Saffran et al., 1999; Thiessen & Saffran, 2009; see Nguyen et al., 2023 for a review). However, the developmental trajectory of such musical sensitivity across the first year of life is still underexplored. Here, we provide a detailed characterization of the progressive maturation of auditory evoked potentials, neurophysiological measures that generally capture rapid changes in the sensory

environment, such as the onset of musical notes (Kushnerenko et al., 2002; Somervail et al., 2021). Specifically, we observed ERPs with progressively shorter latency and larger amplitude throughout the first year of life. While these findings likely reflect a general maturation of sensory systems, they also highlight how the evoked potentials elicited by musical stimuli were enhanced in amplitude compared to those elicited by shuffled (i.e., structure-free) stimuli, an effect consistently observed across all age groups. This is in line with the idea that sensitivity to musical regularities emerges early and persists throughout infancy (Trevarthen, 1999). Indeed, sensitivity to simple rhythmic or pitch regularities has been observed in infants at various ages (Cirelli et al., 2016; Edalati et al., 2023; Flaten et al., 2022; Háden et al., 2009, 2024; Stefanics et al., 2009; Winkler et al., 2009), and, in the case of rhythm, even in premature neonates (Saadatmehr et al., 2025). Such sensitivity may be rooted in general auditory predictive processing, whereby the brain extracts regularities from past observations to generate and update predictions about incoming sensory information (Friston, 2010; Köster et al., 2020; Vuust & Witek, 2014). Ecologically valid music, such as the refrains used here, naturally incorporates rhythmic and pitch regularities that could trigger the generation of predictions. This process was markedly dampened or interrupted by shuffled music (Bianco et al., 2024, 2025; Lense et al., 2022), a finding that could be interpreted as evidence of disengagement from such highly unpredictable sequences (Dayan et al., 2000; Esber & Haselgrove, 2011; Kidd et al., 2014).

Which specific musical regularities drove infants' predictions in our study? Previous EEG research suggests that in newborns, probabilistic auditory predictions are primarily driven by timing rather than pitch regularities when listening to ecologically valid music (Bianco et al., 2025). In contrast, adults rely on both timing and pitch regularities to generate predictions (Di Liberto et al., 2020). Given that our study examined infants older than newborns but not yet as mature as adults, it is likely that timing regularities played a key role for the youngest participants, while pitch regularities may have had a greater influence on the older age groups. However, it is important to note a limitation of such interpretation of our findings: our shuffling procedure permuted both the timing and pitch of the tones simultaneously, which prevents us from disentangling whether the stronger neural responses observed for the music condition reflect sensitivity to rhythmic regularity, melodic structure, or their combination. Hence, to clarify this issue, future research should investigate when, during development, the human brain begins incorporating pitch alongside timing regularities to generate musical predictions.

To explore the underlying neural mechanisms driving the above-described neural process, we analysed ASSR besides ERPs. This was done under the assumption that if the described results were driven by neural entrainment to the beat of the music, then ASSR might better capture differences across conditions, as this measure can capture oscillatory activity. Instead, if the results were driven by evoked responses, then the ASSR results would generally align with ERP results or even be slightly less robust, given that frequency-domain analysis is less suited for capturing time-domain neural modulations. Results showed enhanced ASSR at a frequency matching the musical beat, with stronger power in response to music as opposed to shuffled stimuli and no differences between high- and low-pitch musical stimuli. Hence, ASSR and ERP results were very similar, a conclusion that is in line with previous accounts suggesting that on some occasions, the two measures might originate from the same signal, i.e., evoked responses (Capilla et al., 2011; Novembre & Iannetti, 2018). Further, building on evidence indicating that the infantile P1 originates from the auditory cortex (Chen et al., 2016; Musacchia et al., 2017; Riva et al., 2018), this result sheds light on one of several underlying neural structures that might be responsible for identifying musical structure in the auditory input.

Motor engagement with Music: Movement Sophistication develops with Age

In adults, music engages not only the auditory system but also the motor system (Fujioka et al., 2012; Grahn & Brett, 2007; Novembre & Keller, 2018; Phillips-Silver & Keller, 2012), often leading to spontaneous motor responses coordinated with the musical input (Hurley et al., 2014; Janata et al., 2012). While some evidence indicates that such spontaneous body movements are also exhibited by infants (Provasi et al., 2014), it remains unclear how this behaviour matures throughout development and how sophisticated it is.

First, as a coarse measure of auditory-motor coupling, we used Granger Causality to predict the time course of body movements using the time course of the auditory input (specifically the amplitude envelope). Taking this approach, we could predict body movements of infants from 3 to 12 months, specifically while listening to music as opposed to the shuffled conditions. This result indicates that recognising musical structures leads not only to distinct neural encoding patterns (as discussed in the previous section) but also to different levels of motor engagement. Further, as this result was observed across all ages, we might speculate that such audio-motor coupling, specifically triggered by music, emerges early in development and might be biologically predisposed. This conclusion aligns with previous findings showing that even foetuses as young as 28–35 weeks gestational age

exhibit movement responses to musical sounds (Kisilevsky et al., 2014). However, it is important to note that these prenatal motor responses were not compared to motor responses to other auditory stimuli, as in our study, therefore raising questions about their specificity.

Next, taking a more complex measure of movement, we used principal component analysis to break down body kinematics into 10 independent principal movements (PMs; Bigand et al., 2024; Toiviainen et al., 2010) that explained nearly 80% of the whole kinematic variance. To the best of our knowledge, this approach has never been adopted for the analysis of infants' datasets, which normally do not distinguish between different kinds of movements (Fujii et al., 2014; Nguyen, Reisner, et al., 2023; Zentner & Eerola, 2010) or do so qualitatively (Thelen 1979, 1981). Building on this data-driven approach, we compared to what extent different movements were exhibited by infants across different ages and conditions. We found that only 12-month-old infants exhibited more movement in response to music compared to shuffled music. This result was driven by specific upper-body movements such as front-back rocking, side swaying, proto-clapping, and arm pedalling. It is not straightforward to interpret why these (and not other) movements were triggered by music, and why only at this age. Potential explanations include the development of refined postural control, typically achieved by 9–10 months (Hadders-Algra, 2005), and the fact that infants were tested in a seated position (a measure taken to simplify comparability across age groups). This seating arrangement likely facilitated upper body movements as the feet were resting. Further studies are needed to systematically characterize infants' movements to music across different contexts.

Finally, we examined movement coordination – a potential precursor of dance – and tested whether the periodicity of spontaneous movements matched the periodicity of the music. We did not find any evidence of movement coordination in none of the age groups. This result is in line with previous studies (Fujii et al., 2014; Zentner & Eerola, 2010) and suggests that sensorimotor transformation of music, specifically the ability to preserve its periodicity, develops after the first year of life. This delayed emergence may be due to motor control skills required for such transformation, which indeed are known to continue to mature during toddlerhood and even middle childhood (Kim & Schachner, 2022; Phillips-Silver et al., 2024). Hence, while coarse auditory-motor coupling is present at all ages, diversified movement patterns emerge only by 12 months, and spontaneous coordination with music likely continues to develop throughout infancy into childhood. Taking a broader perspective on infants' motor development, our findings align with research on locomotion across the first 14 months of life, which shows that as the number of motor primitives increases, their

intrinsic variability decreases (Hinnekens et al., 2023). Viewed together, these patterns point toward a gradual refinement of motor control: the human motor system first develops the capacity to control individual muscles, and gradually the capacity to integrate them into motor synergies that support complex, coordinated behaviours, such as locomotion, musical synchronization, and dance.

We suggest that the increasing complexity of infants' motor response to music is linked to the gradual maturation of the dorsal auditory stream, which connects posterior regions of the superior temporal gyrus with premotor cortices (e.g., Chen et al., 2009; Grahn & Brett, 2007; Kotz et al., 2018). Although this pathway is present at birth (Friederici, 2011; Perani et al., 2010), it has been suggested to play a crucial role in rhythmic entrainment and beat perception, functions that likely develop further as the pathway matures (Honing, 2018; Merchant & Honing, 2014; Patel & Iversen, 2014). This developmental trajectory aligns with neuroimaging evidence showing that while the ventral linguistic pathway (connecting temporal and frontal regions via the extreme capsule) is well-established at birth, the dorsal pathway—particularly the arcuate fasciculus connecting temporal regions to inferior frontal areas—continues maturing throughout the first postnatal months, with different maturational timelines for dorsal versus ventral connections (Dubois et al., 2016). Our suggestion is indirectly supported by comparative work on non-human primates, whose dorsal auditory stream is less developed than in humans (Merchant & Honing, 2014; Patel & Iversen, 2014). Indeed, research suggests that adult macaques struggle to recognise musical beats (Honing et al., 2012, 2018), while adult chimpanzees—though capable of adjusting their rhythmic sway periodicity in response to the beat (Hattori, 2021; Hattori & Tomonaga, 2020)—are unable to match it accurately, much like infants in our study and previous literature (Zentner & Eerola, 2010).

Pitch Sensitivity: Auditory and Movement preferences for High-Pitched music

At 6 months, infants exhibited enhanced neural responses to high-pitch music compared to low-pitch music, a specificity not observed at 12 months or in adults. This transient enhancement may reflect a critical period of heightened auditory plasticity, potentially supporting the detection of socially relevant stimuli (Fernald & Kuhl, 1987; Kuhl, 2010). Indeed, high-pitched sounds are prominent in infant-directed speech and singing, which are integral to caregiver-infant interactions and social communication during early infancy (Fernald & Kuhl, 1987; Hilton et al., 2022; Nakata & Trehub, 2011; Trainor & Zacharias, 1998; Tsang & Conrad, 2010). At six months, caregiver-infant face-to-face

interactions peak (Beebe et al., 2016; Feldman, 2007), with infants relying more on behavioural cues, often conveyed through auditory modalities (Gratier et al., 2015; Nguyen, Zimmer, et al., 2023), and less on physical objects than in later developmental stages. Thus, the enhanced sensitivity to high-pitched music at 6 months may reflect an essential phase in infants' developing abilities to process both musical and communicative signals (Shenfield et al., 2003). Toward the end of the first year, attentional focus widens (Cooper et al., 1997; Newman & Hussain, 2006) and face-to-face interactions decrease steeply (Jayaraman et al., 2015). At this stage, strengthening of lower-pitch processing likely leads to more balanced encoding of high- and low-pitch music by 12 months and beyond.

It should also be noted that our musical stimuli comprised polyphonic (two-voice) music, carrying sound frequencies falling within the typical range of infant-directed song (~200-400 Hz, Cirelli et al., 2020; Nguyen, Reisner, et al., 2023; Trainor & Zacharias, 1998). As such, our results might specifically speak for infants' ability to separate (and prioritize among) simultaneous communicative auditory streams (Marie & Trainor, 2013; Trainor, 2015). Indeed, other studies presenting one-voice pure tone sequences (single isochronous and isotonous tones) with high vs. low pitch - notably at frequencies outside our range (130 vs. 1237 Hz) - have reported stronger neural responses to relatively low frequencies (Lenc et al., 2023). Together, these contrasting observations suggest that pitch prioritization changes not only throughout development but also depends on the polyphonic complexity and spectral characteristics of the perceived stimuli. Further research might investigate this interesting issue further.

Moving to the movement results, it remains difficult to explain why high-pitch music was generally associated with enhanced auditory-motor coupling compared to low-pitch music. One possibility is that high-pitch music led to higher arousal, akin to infant-directed songs (Cirelli et al., 2020; Juslin & Laukka, 2003), which in turn strengthened the coupling between spontaneous movements and music. Alternatively, or complementarily, low-pitch music might have led to lower arousal or generally reduced attentional processes, thereby weakening auditory-motor coupling. If increased arousal were to result in greater overall movement, we would expect higher movement levels in the high pitch condition; however, this was not observed. QoM analyses based on the PMs did not reveal significant differences between the high pitch and low pitch conditions. This discrepancy may arise because Granger causality captures subtler temporal coupling between movement and music rather than gross movement quantity. Thus, high-pitch music may modulate the timing and coordination of motor responses without necessarily increasing the overall amount of movement. In line with

prior work (e.g., Bigand et al., 2024), this interpretation emphasizes that musical coordination often involves changes in coupling strength rather than movement quantity per se. Either way, it is difficult to reconcile the movement results with the ERP results showing enhanced auditory processing of high-pitch music only at 6 months. Perhaps, as infants mature, their sensitivity to high-pitched sounds may integrate with not only broader auditory but also motor processing mechanisms. Such a developmental trajectory might explain the reduced neural specificity to high-pitch music by 12 months, alongside a shift toward more coordinated motor responses to lower-pitched stimuli in adulthood (Cameron et al., 2022; Hove et al., 2014; Stupacher et al., 2016). These speculations could be addressed by future research exploring how musical pitch and rhythm interact to drive movement beyond infancy into adulthood to particularly examine *when* neural and motor engagement with low-pitch music becomes more prominent (Cameron et al., 2022; Stupacher et al., 2013, 2016).

Conclusion

Our study demonstrates that while robust auditory processing of music is present as early as 3 months, the translation of these sensory processes into organised motor behaviours unfolds gradually over and beyond the first year of life. Specifically, while coarse auditory-motor coupling is present at all ages, diversified movement patterns emerge only by 12 months, and spontaneous coordination with music likely develops throughout infancy and into childhood. Hence, our study provides evidence that, much like the auditory encoding of music, the propensity to move in response to music emerges early in development. This may reflect a biological or early-developing predisposition, eventually leading to dance-like behaviour. However, by 12 months, these motor responses remain relatively underdeveloped. Additionally, our study points to a previously unknown association between high-pitch music and auditory-motor coupling, extending the current research focus on infant-directed communication towards motor engagement and spontaneous behaviour. Together, these findings provide initial insights into how the developing brain gradually transforms music into spontaneous movements of increasing complexity. Future research should extend our characterisation of music-induced movement beyond the first year of life and further explore its (to-date mysterious) functional significance (Hoehl et al., 2021; Markova et al., 2019; Nguyen, Flaten, et al., 2023; Wass et al., 2020).

Materials & Methods

The study was preregistered: https://aspredicted.org/WW3_3TF. Deviations are listed in the Supplement.

Infant Experiment

Ninety-eight full-term infants, divided into three different age groups (see Fig. 1C), participated in this experiment. Nineteen infants were excluded due to fussiness (n=9) or technical issues (n=10), leading to a 19,6 % attrition rate (in line with infant EEG studies; Hoehl & Wahl, 2012). The remaining 79 infants were either 3 months (n=26, 14 girls, *mean*=113.04 days, *SD*=5.68 days, *range*=98-120 days), 6 months (n=26, 14 girls, *mean*=195.88 days, *SD*=9.46 days, *range*=182-211 days) or 12-13 months of age (n=27, 9 girls, *mean*=380.44 days, *SD*=14.93 days, *range*=361-413 days). All infants were born with a minimum gestational age of 37 weeks, weighed at least 2500 g at birth, and had no known developmental delays, neurological disorders, or hearing impairments. A primary caregiver gave written informed consent and was present during the experiment. The study was approved by the local Ethics committee of the University of Vienna (no. 00645) in line with the Declaration of Helsinki.

Stimuli

Auditory stimuli were generated by rearranging the refrains of two polyphonic children's songs ("La Vaca Lola" and "Hopp Juliska") using Logic Pro X (Apple, Inc.). Stimuli belonged to four distinct conditions (Fig. 1B). The music condition entailed the presentation of the rearranged refrains. The shuffled music condition entailed a disruption of pitch order and timing regularities of the refrains presented in the music condition. Specifically, the original temporal order of the notes was shuffled (disrupting pitch order), and the IOIs were replaced by a new pool of random values uniformly distributed around the original mean IOI ± 30-50% of the distance between the mean and the minimum original IOI (disrupting timing regularities). The bass and melody notes were treated independently in the shuffling process. In addition, the new pool of IOI was not quantized, so the notes did not fall on the beat, thus disrupting the isochrony of the original playsongs. In the high-pitch condition, the melody was shifted one octave higher than in the music condition. In the low-pitch condition, the bassline was shifted one octave lower than in the music condition. Hence, the two voices composing the high-pitch condition were one octave higher than those composing the low-pitch condition. All stimuli were built using the same musical

instruments: the melody was played by a flute ("VHS Flute" from Yamaha DX7, Arturia) and the bassline was played by an electric bass ("Liverpool" electric bass from Logic Pro), with almost all bassline notes (15 out of 16) falling on the beat. Importantly, all stimuli had the same duration (21 s), tempo (135 BPM), tonality (C-major) and loudness (which stayed within a range of 1.5 LUFS, a measure that considers the sensitivity of the human auditory system across frequencies).

Procedure

Setup and Task

Infants were seated in an infant highchair equipped with an age-appropriate baby seat facing a computer screen (1 m distance, see Fig. 1A). Parents sat next to the infants (on their right and slightly towards the back). The stimuli were presented free field (~60 dB SPL) from two audio speakers (Q Acoustics 3020i) placed on either side of the monitor, facing the infants. During the procedure, infants passively listened to a series of auditory stimuli in the following sequence: First, there were 10 seconds of silence, followed by the shuffled music condition. Then, the music, high-pitch and low-pitch conditions were played in random order. After that, there were another 10 seconds of silence, followed by a second round of the music, high-pitch and low-pitch conditions in random order. Finally, the shuffled music condition was played one last time, followed by 10 seconds of silence. The same song was used in all the different conditions within each presentation block. The order of presentation blocks for each song was counterbalanced. Meanwhile, infants were watching a silent movie of blooming flowers, slowly fading in and out (image duration 10 s). Videos were included to keep infants calm and engaged. Only infants who remained calm for at least two presentation blocks were included in the final sample (n=3 were excluded). Stimulus presentation and the synchronization between EEG and video cameras were controlled in Presentation (Version 23.0, Neuro-behavioral System, Berkeley, CA) utilizing triggers sent at the start of the experiment and the beginning of each trial.

Electroencephalography (EEG). We recorded EEG using a Brain Products BrainAmp DC system with 32 active Ag-AgCl electrodes, mounted on infant-sized EEG caps (Acticap) according to the international 10-10 system. The reference electrode was placed at TP9 (left mastoid), and the ground electrode was placed at Fp1. EEG data were sampled at 1000 Hz. Impedances were below 20 kΩ at the beginning of the experiment.

Video Recording. Infant body movement was recorded using three video cameras (Axis Communication, Lund) sampling at 25 Hz with a resolution of 1920 x 1080 pixels. The

three video recordings, taken from frontal (0°), diagonal (45°) and side view (90°) perspectives, were synchronized in VideoSyncPro (Mangold International). The cameras were positioned at the height of the infants' heads to capture their whole bodies.

Adult Control Experiment.

Twenty-six healthy young adults (11 female, *mean* = 27.55 years of age, *SD*: 7.09 years, *range*=19-47 years) participated in the experiment. The study was approved by the Regional Ethics Committee of Liguria (no. 794/2021). Participants underwent the same experimental procedure as the infants, including the same auditory stimuli and video material. EEG was recorded with 64 Ag-AgCl electrodes, closely matched to the infant configuration (see Supplements for further details) and sampled at 1024 Hz. Data were down-sampled to 1000 Hz to match the pre-processing and data analysis procedure (detailed below) of the infant study. The experiment was conducted to provide a ground truth for the interpretation of the infants' EEG data. Therefore, no video data was recorded to track adults' movements.

Data Processing

EEG pre-processing. Infant EEG data are notoriously noisy as infants may not understand or comply with task instructions. It follows that their behaviour can lead to artefacts in the EEG recordings. To mitigate this issue, we used a combination of open-access denoising algorithms and adopted a fully data-driven pre-processing pipeline that we previously developed to denoise EEG data recorded from awake monkeys (Bianco et al., 2024) and humans dancing (Bigand, Bianco, Abalde, Nguyen, et al., 2024). The pipeline relied on a combination of MATLAB functions developed for toolboxes such as Fieldtrip (Oostenveld et al., 2010) and eeglab (Delorme & Makeig, 2004). Continuous EEG data were band-pass filtered (0.3 Hz - 30 Hz, 3rd order Butterworth filter, zero-phase) and segmented into trials starting 3 seconds before song onset and ending 3 seconds post song offset. Next, noisy and faulty electrodes were identified and rejected by assessing flat-lining for over 5 s (function *clean_flatlines*), correlations between electrodes for $r<.1$ or line noise above 20 SD from the mean of all electrodes (function *clean_channels)*. These initial steps were conducted to find a rough indication of electrodes carrying low-quality signals. In addition, mean, standard deviation, and peak-to-peak distances were calculated over time for each electrode. If any of those variables exceeded 2.5 SD from the mean of the other electrodes, that electrode was (provisionally) discarded. This process was iterated without the outlier electrode(s) until a distribution without outliers was found. Electrodes were then re-

referenced to the averaged mastoids (TP9, TP10) or the left mastoid only in cases when TP10 was noisy (N=26). We then used artifact subspace reconstruction (ASR, threshold 5) to remove artefactual EEG activity (Kothe & Jung, 2015). An automatic independent component analysis (ICA) procedure using ICLabel was implemented, and components that were classified as eye components with a minimum probability of 50 % were rejected (mean=0.76, SD=0.67, 0-2 components per infant). Next, previously excluded electrodes were interpolated by using the neighbouring electrodes (spherical spline; mean number of interpolated channels=6, SD=3.09). We assessed data quality (including trial counts and signal-to-noise ratio [SNR]) across age groups and confirmed that these metrics were comparable (see Supplementary Materials for details).

Video data pre-processing. We extracted infants' movement time series by using DeepLabCut (version 2.2.3; Mathis et al., 2018; Nath et al., 2019) in Python on videos filming the infant from the front (see *Fig. 1*). A trained coder (T.N.) labelled 18 body parts per individual (left eye, right eye, mouth, left shoulder, right shoulder, chest, left elbow, left wrist, left hand, right elbow, right wrist, right hand, left knee, left ankle, left foot, right knee, right ankle, right foot) in 10 frames randomly taken from the videos of each infant in each age group (n=24 [3m], 25 [6m] and 24 [12m]; 80% were used for training). We used a ResNet-50 neural network with default parameters for single subject detection for around 5-17 training iterations (correcting labels of 5 frames per video of age groups) until test error plateaued. The test error was 8.27 (3-month-olds), 10.19 (6-month-olds), and 11.89 (12-month-olds) pixels (image size was 1920 by 1080 pixels). The networks of each age group were used to analyse videos from the same age group. The (x and y) coordinates were then pre-processed using a custom code written in Python. Coordinates that were estimated with likelihood values 2 SD away from the mean probability per age group were set to NaN (i.e., not a number). These gaps in the coordinates of body parts were then recovered from predefined neighbouring body parts in a forward and backward approach (from head to feet, followed by feet to head). For example, if the left wrist was missing at t_0, it was recovered using the position of the left elbow or the left hand at t_0. Next, we assessed the correctness of the estimated body part displacements, i.e., demeaned coordinates. Specifically, we set the coordinates (either x or y) of each body part below or above 3 SD from the mean of all trials of each infant to NaN (outlier procedure). To further verify left- and right-hand displacements, we checked the distance between the wrist and hand, specifically using the

same outlier procedure. The remaining NaN values were replaced using linear interpolation for each body part coordinate in time.

Principal movements. We used Principal Component Analysis (PCA) to reduce dimensionality of the pre-processed movement data and at the same time to extract a set of interpretable Principal Movements (PM) that generalize across trials, conditions, and all infants (Bigand, Bianco, Abalde, & Novembre, 2024). To ensure that PCA captured only common movements across participants, rather than postural or anthropometric differences between them, we demeaned and standardized the body part vectors for each trial. Specifically, we subtracted the mean body part vector (averaged over time) from the body part vectors at each frame. Then, these demeaned body part vectors were divided by a global measure of standard deviation across time, which was computed by combining all body parts to preserve the variance differences between body parts. This allowed us to concatenate all trials from all participants into a single data matrix, with each trial contributing equally to the variance of the pooled matrix. PMs were derived from the eigenvectors, and their corresponding scores obtained from the PCA applied to the data matrix:

$$PM(t) = c_i(t)\, w_i$$

where $c_i(t)$ is the ith PM score at time t and w_i is the ith PM weight vector, or eigenvector (1 × 36). Each of the PMs reflects the trajectory of covarying body parts. We extracted 10 PMs that explained 79.7 % of the whole kinematic variance (Fig. 3). It is important to note that the PCA does not directly extract "real movements" but rather movement dimensions with large variances. PMs' time series (score of each principal component) were low-pass filtered at 10 Hz. Next, we calculated the time series' first-degree derivative to extract velocities and the absolute values of velocities to calculate movement quantities (e.g., Bigand et al., 2024; Górecki & Łuczak, 2013). Velocities include information on movement direction (below or above 0) and are, therefore, particularly useful to analyse correlations with other streams, such as music (detailed in the section on Granger Causality). Movement quantities only consider the amount of movement and can contribute to understanding descriptive differences between conditions, regardless of the movement direction.

Data analyses

EEG: Event-related potentials (ERP) analysis. EEG data were first lowpass filtered at 20 Hz (FIR-type filter) to specifically retain the part of the EEG signal carrying event-related potentials. EEG data were epoched according to the onsets of the bass notes of the

musical stimuli (which fell on the beat of the music in 15 out of 16 notes). Epochs included 600 ms of data (-100 ms to +500 ms relative to tone onset – note that notes occurred every 444 ms, so the last 56 ms of each epoch refers to the next tone). In preparation for further analysis, epochs were automatically examined to check if they were still contaminated by artifacts. Epochs were automatically excluded if the standard deviation of the analysed channels exceeded a voltage threshold of 100 µV within a 200 ms interval. All infants and adults contributed at least 165 clean epochs per condition (3m: *M*=340.09, *SD*=138.96; 6m: *M*=373.69, *SD*=144.07; 12m: *M*=361.47, *SD*=157.46; adults: *M*=399.95, *SD*=8.50). Epochs were baseline-corrected by subtracting the average voltage between -100 and 0 ms relative to the tone onset. Epochs belonging to the same experimental condition were averaged together. Event-related EEG amplitude modulations were compared using cluster-based permutation analyses (n_{Perm}=1000; Maris & Ostenveld, 2007). Clusters were based on both temporal consecutiveness and spatial adjacency of EEG electrodes (3 cm distance). A cluster had to be composed of at least two consecutive time points with a *p-value* <.05 on at least two neighbouring EEG electrodes. The clusters were tested using a *p-value* of <.05 and one-tailed according to the above-detailed hypotheses.

In additional analyses, we extracted the amplitude and latencies of the individual peak of the P1 component. Using two linear mixed-effects models, we tested whether amplitudes (dependent variable) and latencies (dependent variable) were dependent on age and condition (fixed and interaction effect) while assuming a random intercept for each participant.

EEG: Auditory Steady State Response (ASSR) analysis. EEG data were segmented according to each condition (21 seconds long, 3 repetitions of the refrain) and averaged over segments of the same condition to enhance the ASSR (see Cirelli et al., 2016). Each segment was submitted to single-taper frequency transformation using a Hilbert taper to estimate frequencies between 0.5 and 5 Hz, with a frequency resolution of 0.25 Hz. Activity unrelated to the frequency of interest was removed by subtracting the average amplitude measured at neighbouring frequency bins (i.e., −5 to −3 bins and +3 to +5 bins). Performing this subtraction removes residual background noise around the frequency bin of interest, leaving only the activity directly related to the ASSR. Next, we extracted amplitudes at the beat-related (2.25 Hz) frequencies and compared these across conditions.

Movement: Granger Causality analysis. We conducted a Granger Causality analysis in R (function = *causality,* library = *vars)* to measure to what extent the envelope of the auditory stimuli predicted movement velocity over time and across conditions (see also Klein et al., 2022). To do so, we extracted the broadband amplitude envelope of the auditory

stimuli using Hilbert transform. We then conducted Granger Causality analyses on amplitude envelopes predicting the PM velocity time series across different model orders (lags) ranging from 40 ms to 1000 ms. Granger Causality is based on the concept of prediction in the sense that if a time series X_t "Granger-causes" another time series Y_t, then past values of X_t contain information that helps predict Y_t beyond the information contained in past values of Y_t alone. Granger Causality is particularly useful for capturing the temporal relationship between different time series that are characterized by sudden and transient bursts in activity (such as infant rhythmic movement; Thelen, 1979). As a sanity check, we also conducted a reverse analysis predicting music (i.e., the amplitude envelope of the auditory stimuli) from the PM velocity time series and verified that the resulting F-Values (averaged across all lags: 40 – 1000 ms) were statistically lower compared to those yielded by the former (music-to-movement) analysis. Indeed, the sanity check analysis reveals that predictions from audio stimuli to PM velocity time series (M=1.08, SE=0.009) were significantly higher than from PMs to audio stimuli (M=1.03, SE=0.009; $\chi^2(1)=1669.2$, $p<.001$, see *Fig. 5*).

Movement: Event-related analysis. To examine whether the events (i.e., musical notes) comprised within the auditory stimuli would trigger event-related movements, we used an event-related approach (mirroring the ERP analysis discussed above). First, we segmented epochs of 700 ms (between -100 ms and +600 ms relative to peaks in the amplitude). Epochs were then averaged for each condition. Detailed results for this analysis are reported in the Supplements (see *Fig. S2*).

Movement: Periodicity analysis. We analysed the movement periodicity using autocorrelation on the same epochs employed in the event-related movement analysis. This means we estimated the autocorrelation of the movement time series in a running window of 3.5 s (length of 1 bar) with 25% overlap, mirroring the approach by Zenter & Eerola (2010). This approach provides an estimation of the periodicity of the PMs. We extracted the local maxima of the autocorrelation values in each window and fitted a normal distribution (Kernel density estimation) to the collection of these values for each condition. We used a bin width of 0.08 and estimated values between 300 and 600 ms according to the beat-related lag in our auditory stimuli. We then compared the beat-related lag (444 ms lag) across conditions. The results are reported in the Supplements (see *Fig. S3*).

Statistical analyses.

Analyses were conducted in R and included non-parametric tests or linear mixed-effects models (LMM; lme4 package). Each participant was modelled as a random intercept in all

LMMs. Statistical significance was evaluated by likelihood-ratio tests conducted using *Anova* (car package), and post-hoc contrasts were run using *emmeans* (emmeans package). We controlled for increased Type I error from multiple comparisons, when necessary, using the false-discovery rate (Benjamini & Hochberg, 1995). The statistical significance level was set to $\alpha = 0.05$.

References

Beebe, B., Messinger, D., Bahrick, L. E., Margolis, A., Buck, K. A., & Chen, H. (2016). A systems view of mother-infant face-to-face communication. *Developmental Psychology*, *52*(4), 556–571. https://doi.org/10.1037/a0040085

Bianco, R., Tóth, B., Bigand, F., Nguyen, T., Sziller, I., Háden, G. P., Winkler, I., & Novembre, G. (2025). *Human newborns form musical predictions based on rhythmic but not melodic structure* (p. 2025.02.19.639016). bioRxiv. https://doi.org/10.1101/2025.02.19.639016

Bianco, R., Zuk, N. J., Bigand, F., Quarta, E., Grasso, S., Arnese, F., Ravignani, A., Battaglia-Mayer, A., & Novembre, G. (2024). Neural encoding of musical expectations in a non-human primate. *Current Biology*, *34*(2), 444-450.e5. https://doi.org/10.1016/j.cub.2023.12.019

Bigand, F., Bianco, R., Abalde, S. F., Nguyen, T., & Novembre, G. (2024). *Imaging the dancing brain: Decoding sensory, motor and social processes during dyadic dance* (p. 2024.12.17.628913). bioRxiv. https://doi.org/10.1101/2024.12.17.628913

Bigand, F., Bianco, R., Abalde, S. F., & Novembre, G. (2024). The geometry of interpersonal synchrony in human dance. *Current Biology*, *34*(13), 3011-3019.e4. https://doi.org/10.1016/j.cub.2024.05.055

Brown, S. (2022). Group dancing as the evolutionary origin of rhythmic entrainment in humans. *New Ideas in Psychology*, *64*, 100902. https://doi.org/10.1016/j.newideapsych.2021.100902

Cameron, D. J., Dotov, D., Flaten, E., Bosnyak, D., Hove, M. J., & Trainor, L. J. (2022). Undetectable very-low frequency sound increases dancing at a live concert. *Current Biology*, *32*(21), R1222–R1223. https://doi.org/10.1016/j.cub.2022.09.035

Capilla, A., Pazo-Alvarez, P., Darriba, Á., Campo, P., & Gross, J. (2011). Steady-State Visual Evoked Potentials Can Be Explained by Temporal Superposition of Transient Event-Related Responses. *PloS One*, *6*, e14543. https://doi.org/10.1371/journal.pone.0014543

Chen, A., Peter, V., & Burnham, D. (2016). Auditory ERP response to successive stimuli in infancy. *PeerJ*, *4*, e1580. https://doi.org/10.7717/peerj.1580

Chen, J. L., Penhune, V. B., & Zatorre, R. J. (2009). The Role of Auditory and Premotor Cortex in Sensorimotor Transformations. *Annals of the New York Academy of Sciences*, *1169*(1), 15–34. https://doi.org/10.1111/j.1749-6632.2009.04556.x

Cirelli, L. K., Jurewicz, Z. B., & Trehub, S. E. (2020). Effects of Maternal Singing Style on Mother–Infant Arousal and Behavior. *Journal of Cognitive Neuroscience*, *32*(7), 1213–1220. https://doi.org/10.1162/jocn_a_01402

Cirelli, L. K., Spinelli, C., Nozaradan, S., & Trainor, L. J. (2016). Measuring Neural Entrainment to Beat and Meter in Infants: Effects of Music Background. *Frontiers in Neuroscience*, *10*. https://doi.org/10.3389/fnins.2016.00229

Conrad, N. J., Walsh, J., Allen, J. M., & Tsang, C. D. (2011). Examining infants' preferences for tempo in lullabies and playsongs. *Canadian Journal of Experimental Psychology / Revue Canadienne de Psychologie Expérimentale*, *65*(3), 168–172. https://doi.org/10.1037/a0023296

Cooper, R. P., Abraham, J., Berman, S., & Staska, M. (1997). The development of infants' preference for motherese. *Infant Behavior and Development*, *20*(4), 477–488. https://doi.org/10.1016/S0163-6383(97)90037-0

Costa-Giomi, E., & Sun, X. (2016). Infants' Home Soundscape: A Day in the Life of a Family. In *Contemporary Research in Music Learning Across the Lifespan* (pp. 87–96). Routledge.

Damsma, A., Roo, M., Doelling, K., Bazin, P.-L., & Bouwer, F. (2024). *Tempo-dependent selective enhancement of neural responses at the beat frequency can be explained by both an oscillator and an evoked model*. https://doi.org/10.1101/2024.07.11.603023

Dayan, P., Kakade, S., & Montague, P. R. (2000). Learning and selective attention. *Nature Neuroscience*, *3*(S11), 1218–1223. https://doi.org/10.1038/81504

Delorme, A., & Makeig, S. (2004). EEGLAB: An open source toolbox for analysis of single-trial EEG dynamics including independent component analysis. *Journal of Neuroscience Methods*, *134*(1), 9–21. https://doi.org/10.1016/j.jneumeth.2003.10.009

Di Liberto, G. M., Pelofi, C., Bianco, R., Patel, P., Mehta, A. D., Herrero, J. L., de Cheveigné, A., Shamma, S., & Mesgarani, N. (2020). Cortical encoding of melodic expectations in human temporal cortex. *eLife*, *9*, e51784. https://doi.org/10.7554/eLife.51784

Dubois, J., Poupon, C., Thirion, B., Simonnet, H., Kulikova, S., Leroy, F., Hertz-Pannier, L., & Dehaene-Lambertz, G. (2016). Exploring the Early Organization and Maturation of Linguistic Pathways in the Human Infant Brain. *Cerebral Cortex*, *26*(5), 2283–2298. https://doi.org/10.1093/cercor/bhv082

Eckerdal, P., & Merker, B. (2009). 'Music' and the 'action song' in infant development: An interpretation. In *Communicative musicality: Exploring the basis of human companionship* (pp. 241–262). Oxford University Press.

Edalati, M., Wallois, F., Safaie, J., Ghostine, G., Kongolo, G., Trainor, L. J., & Moghimi, S. (2023). Rhythm in the premature neonate brain: Very early processing of auditory beat and meter. *The Journal of Neuroscience*, JN-RM-1100-22. https://doi.org/10.1523/JNEUROSCI.1100-22.2023

Esber, G. R., & Haselgrove, M. (2011). Reconciling the influence of predictiveness and uncertainty on stimulus salience: A model of attention in associative learning. *Proceedings of the Royal Society B: Biological Sciences*, *278*(1718), 2553–2561. https://doi.org/10.1098/rspb.2011.0836

Feldman, R. (2007). Parent–Infant Synchrony: Biological Foundations and Developmental Outcomes. *Current Directions in Psychological Science*, *16*(6), 340–345. https://doi.org/10.1111/j.1467-8721.2007.00532.x

Fernald, A., & Kuhl, P. (1987). Acoustic determinants of infant preference for motherese speech. *Infant Behavior and Development*, *10*(3), 279–293. https://doi.org/10.1016/0163-6383(87)90017-8

Fernald, A., & Simon, T. (1984). Expanded intonation contours in mothers' speech to newborns. *Developmental Psychology*, *20*(1), 104–113. https://doi.org/10.1037/0012-1649.20.1.104

Fitch, W. T. (2015). Four principles of bio-musicology. *Philosophical Transactions of the Royal Society B: Biological Sciences*, *370*(1664), 20140091. https://doi.org/10.1098/rstb.2014.0091

Flaten, E., Marshall, S. A., Dittrich, A., & Trainor, L. J. (2022). Evidence for top‐down metre perception in infancy as shown by primed neural responses to an ambiguous rhythm. *European Journal of Neuroscience*, ejn.15671. https://doi.org/10.1111/ejn.15671

Friederici, A. D. (2011). The Brain Basis of Language Processing: From Structure to Function. *Physiological Reviews*, *91*(4), 1357–1392. https://doi.org/10.1152/physrev.00006.2011

Friston, K. (2010). The free-energy principle: A unified brain theory? *Nature Reviews Neuroscience*, *11*(2), Article 2. https://doi.org/10.1038/nrn2787

Fujii, S., Watanabe, H., Oohashi, H., Hirashima, M., Nozaki, D., & Taga, G. (2014). Precursors of Dancing and Singing to Music in Three- to Four-Months-Old Infants. *PLOS ONE*, *9*(5), e97680. https://doi.org/10.1371/journal.pone.0097680

Fujioka, T., Trainor, L. J., Large, E. W., & Ross, B. (2012). Internalized Timing of Isochronous Sounds Is Represented in Neuromagnetic Beta Oscillations. *The Journal of Neuroscience*, *32*(5), 1791–1802. https://doi.org/10.1523/JNEUROSCI.4107-11.2012

Górecki, T., & Łuczak, M. (2013). Using derivatives in time series classification. *Data Mining and Knowledge Discovery*, *26*, 310–331. https://doi.org/10.1007/s10618-012-0251-4

Grahn, J. A., & Brett, M. (2007). Rhythm and Beat Perception in Motor Areas of the Brain. *Journal of Cognitive Neuroscience*, *19*(5), 893–906. https://doi.org/10.1162/jocn.2007.19.5.893

Gratier, M., Devouche, E., Guellai, B., Infanti, R., Yilmaz, E., & Parlato-Oliveira, E. (2015). Early development of turn-taking in vocal interaction between mothers and infants. *Frontiers in Psychology*, *6*, 1167. https://doi.org/10.3389/fpsyg.2015.01167

Hadders-Algra, M. (2005). Development of Postural Control During the First 18 Months of Life. *Neural Plasticity*, *12*(2–3), 695071. https://doi.org/10.1155/NP.2005.99

Háden, G. P., Bouwer, F. L., Honing, H., & Winkler, I. (2022). *Beat processing in newborn infants cannot be explained by statistical learning based on transition probabilities* [Preprint]. Neuroscience. https://doi.org/10.1101/2022.12.20.521245

Háden, G. P., Németh, R., Török, M., & Winkler, I. (2015). Predictive processing of pitch trends in newborn infants. *Brain Research*, *1626*, 14–20. https://doi.org/10.1016/j.brainres.2015.02.048

Háden, G. P., Stefanics, G., Vestergaard, M. D., Denham, S. L., Sziller, I., & Winkler, I. (2009). Timbre-independent extraction of pitch in newborn infants. *Psychophysiology*, *46*(1), 69–74. https://doi.org/10.1111/j.1469-8986.2008.00749.x

Hattori, Y. (2021). Behavioral Coordination and Synchronization in Non-human Primates. In J. R. Anderson & H. Kuroshima (Eds), *Comparative Cognition: Commonalities and Diversity* (pp. 139–151). Springer. https://doi.org/10.1007/978-981-16-2028-7_9

Hattori, Y., & Tomonaga, M. (2020). Rhythmic swaying induced by sound in chimpanzees (Pan troglodytes). *Proceedings of the National Academy of Sciences*, *117*(2), 936–942. https://doi.org/10.1073/pnas.1910318116

Hilton, C. B., Moser, C. J., Bertolo, M., Lee-Rubin, H., Amir, D., Bainbridge, C. M., Simson, J., Knox, D., Glowacki, L., Alemu, E., Galbarczyk, A., Jasienska, G., Ross, C. T., Neff, M. B., Martin, A., Cirelli, L. K., Trehub, S. E., Song, J., Kim, M., … Mehr, S. A. (2022). Acoustic regularities in infant-directed speech and song across cultures. *Nature Human Behaviour*. https://doi.org/10.1038/s41562-022-01410-x

Hinnekens, E., Barbu-Roth, M., Do, M.-C., Berret, B., & Teulier, C. (2023). Generating variability from motor primitives during infant locomotor development. *eLife*, *12*, e87463. https://doi.org/10.7554/eLife.87463

Hoehl, S., Fairhurst, M., & Schirmer, A. (2021). Interactional synchrony: Signals, mechanisms and benefits. *Social Cognitive and Affective Neuroscience*, *16*(1–2), 5–18. https://doi.org/10.1093/scan/nsaa024

Hoehl, S., & Wahl, S. (2012). Recording infant ERP data for cognitive research. *Developmental Neuropsychology*, *37*(3), 187–209. https://doi.org/10.1080/87565641.2011.627958

Honing, H. (2018). *The Origins of Musicality*. MIT Press.

Honing, H., Bouwer, F. L., Prado, L., & Merchant, H. (2018). Rhesus Monkeys (Macaca mulatta) Sense Isochrony in Rhythm, but Not the Beat: Additional Support for the Gradual Audiomotor Evolution Hypothesis. *Frontiers in Neuroscience*, *12*. https://doi.org/10.3389/fnins.2018.00475

Honing, H., Merchant, H., Háden, G. P., Prado, L., & Bartolo, R. (2012). Rhesus Monkeys (Macaca mulatta) Detect Rhythmic Groups in Music, but Not the Beat. *PLoS ONE*, *7*(12), e51369. https://doi.org/10.1371/journal.pone.0051369

Honing, H., ten Cate, C., Peretz, I., & Trehub, S. E. (2015). Without it no music: Cognition, biology and evolution of musicality. *Philosophical Transactions of the Royal Society of London. Series B, Biological Sciences*, *370*(1664), 20140088. https://doi.org/10.1098/rstb.2014.0088

Hove, M. J., Marie, C., Bruce, I. C., & Trainor, L. J. (2014). Superior time perception for lower musical pitch explains why bass-ranged instruments lay down musical rhythms. *Proceedings of the National Academy of Sciences*, *111*(28), 10383–10388. https://doi.org/10.1073/pnas.1402039111

Hurley, B. K., Martens, P. A., & Janata, P. (2014). Spontaneous sensorimotor coupling with multipart music. *Journal of Experimental Psychology: Human Perception and Performance*, *40*, 1679–1696. https://doi.org/10.1037/a0037154

Janata, P., Tomic, S. T., & Haberman, J. M. (2012). Sensorimotor coupling in music and the psychology of the groove. *Journal of Experimental Psychology: General*, *141*(1), 54–75. https://doi.org/10.1037/a0024208

Jayaraman, S., Fausey, C. M., & Smith, L. B. (2015). The Faces in Infant-Perspective Scenes Change over the First Year of Life. *PLOS ONE*, *10*(5), e0123780. https://doi.org/10.1371/journal.pone.0123780

Juslin, P. N., & Laukka, P. (2003). Communication of emotions in vocal expression and music performance: Different channels, same code? *Psychological Bulletin*, *129*(5), 770–814. https://doi.org/10.1037/0033-2909.129.5.770

Kidd, C., Piantadosi, S. T., & Aslin, R. N. (2014). The Goldilocks Effect in Infant Auditory Attention. *Child Development*, *85*(5), 1795–1804. https://doi.org/10.1111/cdev.12263

Kim, M., & Schachner, A. (2022). The origins of dance: Characterizing the development of infants' earliest dance behavior. *Developmental Psychology*. https://doi.org/10.1037/dev0001436

Kisilevsky, B. s., Hains, S. m. j., Jacquet, A.-Y., Granier-Deferre, C., & Lecanuet, J. p. (2004). Maturation of fetal responses to music. *Developmental Science*, *7*(5), 550–559. https://doi.org/10.1111/j.1467-7687.2004.00379.x

Klein, L., Wood, E. A., Bosnyak, D., & Trainor, L. J. (2022). Follow the sound of my violin: Granger causality reflects information flow in sound. *Frontiers in Human Neuroscience*, *16*, 982177. https://doi.org/10.3389/fnhum.2022.982177

Köster, M., Kayhan, E., Langeloh, M., & Hoehl, S. (2020). Making Sense of the World: Infant Learning From a Predictive Processing Perspective. *Perspectives on Psychological Science*, *15*(3), 562–571. https://doi.org/10.1177/1745691619895071

Kothe, C. A. E., & Jung, T.-P. (2015). *Artifact removal techniques with signal reconstruction* (World Intellectual Property Organization Patent No. 2015/047462 A9). https://patents.google.com/patent/WO2015047462A9/en

Kotz, S. A., Ravignani, A., & Fitch, W. T. (2018). The Evolution of Rhythm Processing. *Trends in Cognitive Sciences*, *22*(10), 896–910. https://doi.org/10.1016/j.tics.2018.08.002

Kouider, S., Long, B., Le Stanc, L., Charron, S., Fievet, A.-C., Barbosa, L. S., & Gelskov, S. V. (2015). Neural dynamics of prediction and surprise in infants. *Nature Communications*, *6*(1), 8537. https://doi.org/10.1038/ncomms9537

Kuhl, P. K. (2010). Brain Mechanisms in Early Language Acquisition. *Neuron*, *67*(5), 713–727. https://doi.org/10.1016/j.neuron.2010.08.038

Kuhl, P. K., & Meltzoff, A. N. (1982). The bimodal perception of speech in infancy. *Science (New York, N.Y.)*, *218*(4577), 1138–1141. https://doi.org/10.1126/science.7146899

Kushnerenko, E., Ceponiene, R., Balan, P., Fellman, V., Huotilaine, M., & Näätäne, R. (2002). Maturation of the auditory event-related potentials during the first year of life. *Neuroreport*, *13*(1), 47–51. https://doi.org/10.1097/00001756-200201210-00014

Lenc, T., Peter, V., Hooper, C., Keller, P. E., Burnham, D., & Nozaradan, S. (2023). Infants show enhanced neural responses to musical meter frequencies beyond low-level features. *Developmental Science*, *26*(5), e13353. https://doi.org/10.1111/desc.13353

Lense, M. D., Shultz, S., Astésano, C., & Jones, W. (2022). Music of infant-directed singing entrains infants' social visual behavior. *Proceedings of the National Academy of Sciences*, *119*(45), e2116967119. https://doi.org/10.1073/pnas.2116967119

Lijffijt, M., Moeller, F. G., Boutros, N. N., Burroughs, S., Lane, S. D., Steinberg, J. L., & Swann, A. C. (2009). The Role of Age, Gender, Education, and Intelligence in P50, N100, and P200 Auditory Sensory Gating. *Journal of Psychophysiology*, *23*(2), 52–62. https://doi.org/10.1027/0269-8803.23.2.52

Marie, C., & Trainor, L. J. (2013). Development of Simultaneous Pitch Encoding: Infants Show a High Voice Superiority Effect. *Cerebral Cortex*, *23*(3), 660–669. https://doi.org/10.1093/cercor/bhs050

Marie, C., & Trainor, L. J. (2014). Early development of polyphonic sound encoding and the high voice superiority effect. *Neuropsychologia*, *57*, 50–58. https://doi.org/10.1016/j.neuropsychologia.2014.02.023

Markova, G., Nguyen, T., & Hoehl, S. (2019). Neurobehavioral Interpersonal Synchrony in Early Development: The Role of Interactional Rhythms. *Frontiers in Psychology*, *10*, 2078. https://doi.org/10.3389/fpsyg.2019.02078

Mathis, A., Mamidanna, P., Cury, K. M., Abe, T., Murthy, V. N., Mathis, M. W., & Bethge, M. (2018). DeepLabCut: Markerless pose estimation of user-defined body parts with deep learning. *Nature Neuroscience*, *21*(9), Article 9. https://doi.org/10.1038/s41593-018-0209-y

Merchant, H., & Honing, H. (2014). Are non-human primates capable of rhythmic entrainment? Evidence for the gradual audiomotor evolution hypothesis. *Frontiers in Neuroscience*, *7*, 274. https://doi.org/10.3389/fnins.2013.00274

Musacchia, G., Ortiz-Mantilla, S., Choudhury, N., Realpe-Bonilla, T., Roesler, C., & Benasich, A. A. (2017). Active auditory experience in infancy promotes brain plasticity in Theta and Gamma oscillations. *Developmental Cognitive Neuroscience*, *26*, 9–19. https://doi.org/10.1016/j.dcn.2017.04.004

Nakata, T., & Trehub, S. E. (2011). Expressive timing and dynamics in infant-directed and non-infant-directed singing. *Psychomusicology: Music, Mind and Brain*, *21*(1–2), 45–53. https://doi.org/10.1037/h0094003

Nath, T., Mathis, A., Chen, A. C., Patel, A., Bethge, M., & Mathis, M. W. (2019). Using DeepLabCut for 3D markerless pose estimation across species and behaviors. *Nature Protocols*, *14*(7), 2152–2176. https://doi.org/10.1038/s41596-019-0176-0

Nave, K. M., Hannon, E. E., & Snyder, J. S. (2022). Steady state-evoked potentials of subjective beat perception in musical rhythms. *Psychophysiology*, *59*(2), e13963. https://doi.org/10.1111/psyp.13963

Newman, R. S., & Hussain, I. (2006). Changes in Preference for Infant-Directed Speech in Low and Moderate Noise by 4.5- to 13-Month-Olds. *Infancy*, *10*(1), 61–76. https://doi.org/10.1207/s15327078in1001_4

Nguyen, T., Flaten, E., Trainor, L. J., & Novembre, G. (2023). *Early social communication through music: State of the art and future perspectives* [Preprint]. PsyArXiv. https://doi.org/10.31234/osf.io/j5g69

Nguyen, T., Reisner, S., Lueger, A., Wass, S. V., Hoehl, S., & Markova, G. (2023a). Sing to me, baby: Infants show neural tracking and rhythmic movements to live and dynamic maternal singing. *Developmental Cognitive Neuroscience*, *64*, 101313. https://doi.org/10.1016/j.dcn.2023.101313

Nguyen, T., Reisner, S., Lueger, A., Wass, S. V., Hoehl, S., & Markova, G. (2023b). *Sing to me, baby: Infants show neural tracking and rhythmic movements to live and dynamic maternal singing* (p. 2023.02.28.530310). bioRxiv. https://doi.org/10.1101/2023.02.28.530310

Nguyen, T., Zimmer, L., & Hoehl, S. (2023). Your turn, my turn. Neural synchrony in mother–infant proto-conversation. *Philosophical Transactions of the Royal Society B: Biological Sciences*, *378*(1875), 20210488. https://doi.org/10.1098/rstb.2021.0488

Novembre, G., & Iannetti, G. D. (2018). Tagging the musical beat: Neural entrainment or event-related potentials? *Proceedings of the National Academy of Sciences*, *115*(47), E11002–E11003. https://doi.org/10.1073/pnas.1815311115

Novembre, G., & Keller, P. E. (2018). Music and Action. In R. Bader (Ed.), *Springer Handbook of Systematic Musicology* (pp. 523–537). Springer. https://doi.org/10.1007/978-3-662-55004-5_28

Novembre, G., Pawar, V. M., Bufacchi, R. J., Kilintari, M., Srinivasan, M., Rothwell, J. C., Haggard, P., & Iannetti, G. D. (2018). Saliency Detection as a Reactive Process: Unexpected Sensory Events Evoke Corticomuscular Coupling. *The Journal of Neuroscience*, *38*(9), 2385–2397. https://doi.org/10.1523/JNEUROSCI.2474-17.2017

Oostenveld, R., Fries, P., Maris, E., & Schoffelen, J.-M. (2010). FieldTrip: Open Source Software for Advanced Analysis of MEG, EEG, and Invasive Electrophysiological Data. *Computational Intelligence and Neuroscience*, *2011*, e156869. https://doi.org/10.1155/2011/156869

Patel, A. D., & Iversen, J. R. (2014). The evolutionary neuroscience of musical beat perception: The Action Simulation for Auditory Prediction (ASAP) hypothesis. *Frontiers in Systems Neuroscience*, *8*. https://www.frontiersin.org/articles/10.3389/fnsys.2014.00057

Perani, D., Saccuman, M. C., Scifo, P., Spada, D., Andreolli, G., Rovelli, R., Baldoli, C., & Koelsch, S. (2010). Functional specializations for music processing in the human newborn brain. *Proceedings of the National Academy of Sciences*, *107*(10), 4758–4763. https://doi.org/10.1073/pnas.0909074107

Phillips-Silver, J., Hartmann, M., Fernández-García, L., Maurno, N. C. G., Toiviainen, P., & González, M. T. D. (2024). Development of full-body rhythmic synchronization in middle childhood. *Scientific Reports*, *14*(1), 15741. https://doi.org/10.1038/s41598-024-66438-7

Phillips-Silver, J., & Keller, P. (2012). Searching for Roots of Entrainment and Joint Action in Early Musical Interactions. *Frontiers in Human Neuroscience*, *6*. https://doi.org/10.3389/fnhum.2012.00026

Pratt, H., Starr, A., Michalewski, H. J., Bleich, N., & Mittelman, N. (2008). The auditory P50 component to onset and offset of sound. *Clinical Neurophysiology*, *119*(2), 376–387. https://doi.org/10.1016/j.clinph.2007.10.016

Provasi, J., Anderson, D. I., & Barbu-Roth, M. (2014). Rhythm perception, production, and synchronization during the perinatal period. *Frontiers in Psychology*, *5*, 1048.

Remijn, G. B., Hasuo, E., Fujihira, H., & Morimoto, S. (2014). An introduction to the measurement of auditory event-related potentials (ERPs). *Acoustical Science and Technology*, *35*(5), 229–242. https://doi.org/10.1250/ast.35.229

Riva, V., Cantiani, C., Mornati, G., Gallo, M., Villa, L., Mani, E., Saviozzi, I., Marino, C., & Molteni, M. (2018). Distinct ERP profiles for auditory processing in infants at-risk for autism and language impairment. *Scientific Reports*, *8*(1), 715. https://doi.org/10.1038/s41598-017-19009-y

Saadatmehr, B., Edalati, M., Wallois, F., Ghostine, G., Kongolo, G., Flaten, E., Tillmann, B., Trainor, L., & Moghimi, S. (2025). Auditory Rhythm Encoding during the Last Trimester of Human Gestation: From Tracking the Basic Beat to Tracking Hierarchical Nested Temporal Structures. *Journal of Neuroscience*, *45*(4). https://doi.org/10.1523/JNEUROSCI.0398-24.2024

Saffran, J. R., Johnson, E. K., Aslin, R. N., & Newport, E. L. (1999). Statistical learning of tone sequences by human infants and adults. *Cognition*, *70*(1), 27–52. https://doi.org/10.1016/S0010-0277(98)00075-4

Schachner, A., Brady, T. F., Pepperberg, I. M., & Hauser, M. D. (2009). Spontaneous Motor Entrainment to Music in Multiple Vocal Mimicking Species. *Current Biology*, *19*(10), 831–836. https://doi.org/10.1016/j.cub.2009.03.061

Somervail, R., Zhang, F., Novembre, G., Bufacchi, R. J., Guo, Y., Crepaldi, M., Hu, L., & Iannetti, G. D. (2021). Waves of Change: Brain Sensitivity to Differential, not Absolute, Stimulus Intensity is Conserved Across Humans and Rats. *Cerebral Cortex (New York, N.Y.: 1991)*, *31*(2), 949–960. https://doi.org/10.1093/cercor/bhaa267

Stefanics, G., Háden, G. P., Sziller, I., Balázs, L., Beke, A., & Winkler, I. (2009). Newborn infants process pitch intervals. *Clinical Neurophysiology*, *120*(2), 304–308. https://doi.org/10.1016/j.clinph.2008.11.020

Stupacher, J., Hove, M. J., & Janata, P. (2016). Audio Features Underlying Perceived Groove and Sensorimotor Synchronization in Music. *Music Perception*, *33*(5), 571–589. https://doi.org/10.1525/mp.2016.33.5.571

Stupacher, J., Hove, M. J., Novembre, G., Schütz-Bosbach, S., & Keller, P. E. (2013). Musical groove modulates motor cortex excitability: A TMS investigation. *Brain and Cognition*, *82*(2), 127–136. https://doi.org/10.1016/j.bandc.2013.03.003

Thelen, E. (1979). Rhythmical stereotypies in normal human infants. *Animal Behaviour*, *27*, 699–715. https://doi.org/10.1016/0003-3472(79)90006-X

Thelen, E. (1981). Kicking, rocking, and waving: Contextual analysis of rhythmical stereotypies in normal human infants. *Animal Behaviour*, *29*(1), 3–11. https://doi.org/10.1016/s0003-3472(81)80146-7

Thiessen, E. D., & Saffran, J. R. (2009). How the Melody Facilitates the Message and Vice Versa in Infant Learning and Memory. *Annals of the New York Academy of Sciences*, *1169*(1), 225–233. https://doi.org/10.1111/j.1749-6632.2009.04547.x

Toiviainen, P., Luck, G., & Thompson, M. R. (2010). Embodied Meter: Hierarchical Eigenmodes in Music-Induced Movement. *Music Perception*, *28*(1), 59–70. https://doi.org/10.1525/mp.2010.28.1.59

Trainor, L. J. (1996). Infant preferences for infant-directed versus noninfant-directed playsongs and lullabies. *Infant Behavior and Development*, *19*(1), 83–92. https://doi.org/10.1016/S0163-6383(96)90046-6

Trainor, L. J. (2015). The origins of music in auditory scene analysis and the roles of evolution and culture in musical creation. *Philosophical Transactions of the Royal Society B: Biological Sciences*, *370*(1664), 20140089. https://doi.org/10.1098/rstb.2014.0089

Trainor, L. J., & Desjardins, R. N. (2002). Pitch characteristics of infant-directed speech affect infants' ability to discriminate vowels. *Psychonomic Bulletin & Review*, *9*(2), 335–340. https://doi.org/10.3758/BF03196290

Trainor, L. J., & Zacharias, C. A. (1998). Infants prefer higher-pitched singing. *Infant Behavior and Development*, *21*(4), 799–805. https://doi.org/10.1016/S0163-6383(98)90047-9

Trehub, S. E. (2003). The developmental origins of musicality. *Nature Neuroscience*, *6*(7), 669–673. https://doi.org/10.1038/nn1084

Trehub, S. E., Becker, J., & Morley, I. (2015). Cross-cultural perspectives on music and musicality. *Philosophical Transactions of the Royal Society B: Biological Sciences*, *370*(1664), 20140096. https://doi.org/10.1098/rstb.2014.0096

Trevarthen, C. (1999). Musicality and the intrinsic motive pulse: Evidence from human psychobiology and infant communication. *Musicae Scientiae*, *3*(1_suppl), 155–215. https://doi.org/10.1177/10298649000030S109

Tsang, C. D., & Conrad, N. J. (2010). Does the message matter? The effect of song type on infants' pitch preferences for lullabies and playsongs. *Infant Behavior and Development*, *33*(1), 96–100. https://doi.org/10.1016/j.infbeh.2009.11.006

Van Dyck, E., Moelants, D., Demey, M., Deweppe, A., Coussement, P., & Leman, M. (2013). The Impact of the Bass Drum on Human Dance Movement. *Music Perception*, *30*(4), 349–359. https://doi.org/10.1525/mp.2013.30.4.349

Vuust, P., & Witek, M. (2014). Rhythmic complexity and predictive coding: A novel approach to modeling rhythm and meter perception in music. *Frontiers in Psychology*, *5*, 1111. https://doi.org/10.3389/fpsyg.2014.01111

Wass, S. V., Whitehorn, M., Marriott Haresign, I., Phillips, E., & Leong, V. (2020). Interpersonal Neural Entrainment during Early Social Interaction. *Trends in Cognitive Sciences*, *24*(4), 329–342. https://doi.org/10.1016/j.tics.2020.01.006

Winkler, I., Háden, G. P., Ladinig, O., Sziller, I., & Honing, H. (2009). Newborn infants detect the beat in music. *Proceedings of the National Academy of Sciences*, *106*(7), 2468–2471. https://doi.org/10.1073/pnas.0809035106

Wunderlich, J. L., Cone-Wesson, B. K., & Shepherd, R. K. (2006). Maturation of the cortical auditory evoked potential in infants and young children. *Hearing Research*. https://doi.org/10.1016/j.heares.2005.11.010

Zentner, M. R., & Eerola, T. (2010). Rhythmic engagement with music in infancy. *Proceedings of the National Academy of Sciences*, *107*(13), 5768–5773. https://doi.org/10.1073/pnas.1000121107