Individual differences in neural markers of beat processing relate to spoken grammar skills in six-year-old children

Valentina Persici [a,b,c,d,*], Scott D. Blain [e], John R. Iversen [f,g], Alexandra P. Key [h,i,a], Sonja A. Kotz [j,k], J. Devin McAuley [l], Reyna L. Gordon [c,m,a,i,*]

[a] *Vanderbilt Brain Institute, Vanderbilt University, Nashville, TN, USA*
[b] *Department of Psychology, University of Milano – Bicocca, Milan, Italy*
[c] *Department of Otolaryngology – Head & Neck Surgery, Vanderbilt University Medical Center, Nashville, TN, USA*
[d] *Department of Human Sciences, University of Verona, Verona, Italy*
[e] *Department of Psychiatry, University of Michigan, Ann Arbor, MI, USA*
[f] *Department of Psychology, Neuroscience and Behaviour, McMaster University, Hamilton, Ontario, Canada*
[g] *Institute for Neural Computation, University of California San Diego, La Jolla, CA, USA*
[h] *Department of Hearing and Speech Sciences, Vanderbilt University Medical Center, Nashville, TN, USA*
[i] *Vanderbilt Kennedy Center, Vanderbilt University Medical Center, Nashville, TN, USA*
[j] *Department of Neuropsychology and Psychopharmacology, Maastricht University, Maastricht, the Netherlands*
[k] *Department of Neuropsychology, Max Planck Institute for Human Cognitive and Brain Sciences, Leipzig, Germany*
[l] *Department of Psychology, Michigan State University, East Lansing, MI, USA*
[m] *Department of Psychology, Vanderbilt University, Nashville, TN, USA*

ARTICLE INFO

ABSTRACT

Based on the idea that neural entrainment establishes regular attentional fluctuations that facilitate hierarchical processing in both music and language, we hypothesized that individual differences in syntactic (grammatical) skills will be partly explained by patterns of neural responses to musical rhythm. To test this hypothesis, we recorded neural activity using electroencephalography (EEG) while children ($N = 25$) listened passively to rhythmic patterns that induced different beat percepts. Analysis of evoked beta and gamma activity revealed that individual differences in the magnitude of neural responses to rhythm explained variance in six-year-olds' expressive grammar abilities, beyond and complementarily to their performance in a behavioral rhythm perception task. These results reinforce the idea that mechanisms of neural beat entrainment may be a shared neural resource supporting hierarchical processing across music and language and suggest a relevant marker of the relationship between rhythm processing and grammar abilities in elementary-school-age children, previously observed only behaviorally.

1. Introduction

1.1. Commonalities between rhythm and grammar: features, evolutionary basis, and underlying mechanisms

Recent studies have shown commonalities between musical rhythm and grammar. Children who demonstrate better rhythmic perception skills tend to show enhanced performance in syntactic tasks (Gordon, Shivers, et al., 2015; Nitin et al., 2023; Politimou et al., 2019; Swaminathan & Schellenberg, 2019; cf. Ozernov-Palchik et al., 2018),

especially when complex syntactic structures are involved (Gordon, Jacobs, et al., 2015; Lee et al., 2020). Moreover, rhythmic structure affects the processing of grammatical structure in speech in adults (Kotz et al., 2009; Schmidt-Kassow & Kotz, 2008). Rhythmic priming paradigms have also shown that grammar performance is positively affected by brief exposure to metrically regular rhythmic musical primes (Canette et al., 2019; Chern et al., 2018; Ladányi et al., 2021; Przybylski et al., 2013). Such relations between rhythm and grammar have been explained as the result of shared cognitive processes that are recruited when hierarchical structures are processed in both language and music

(Fitch, 2013; Fitch & Martins, 2014; Ladányi, Persici, et al., 2020). Both rhythm and grammar are organized in tree-like structures in which lower levels are incorporated into higher levels and elements are ordered according to specific rules (Fitch, 2017). In music, listeners generally utilize regularly timed basic units (whether accented or not) to extract the underlying pulse, or beat, which serves as an anchor "around which other [rhythmic] events are organized" (Iversen, Repp, & Patel, 2009:58). Beats, in turn, are perceptually grouped and organized into a hierarchical structure called 'meter' (Lerdahl & Jackendoff, 1983), which allows to interpret the music input as patterns of strong and weak beats. Language is also constituted of elements that are ordered following a hierarchical structure and specific rules (Lashley, 1951), with words combining to form phrases and smaller phrases combining and/or moving to form larger phrases. Thus, grammar perception and production also require perception and manipulation of stimuli that unfold serially but that are organized hierarchically.

The ability to perform these operations in rhythm may have an evolutionary basis. Animal studies have shown that perceptually extracting the beat is an ability shared by birds and mammals (Kotz et al., 2018). The pervasive nature of music across cultures (Mehr et al., 2019; Ravignani et al., 2017; Savage et al., 2015) and the finding that engaging in rhythmic activities enhances prosocial behavior and social interaction and bonding (Cirelli et al., 2014; Lang et al., 2017; Pearce et al., 2015; Tarr et al., 2016) have led to the hypothesis that the ability to perceive and synchronize with rhythms may have evolved in humans to facilitate communication and foster group cohesion (Honing, 2018; Patel & Iversen, 2014; Savage et al., 2020). Developmentally, it has been proposed that rhythmic processing guides speech segmentation and comprehension and grammar learning in children (Beier & Ferreira, 2018; Dilley & McAuley, 2008; Kotz et al., 2018; Kotz & Schwartze, 2010, 2016). Research has shown that sensitivity to rhythmic cues in speech such as stress patterns is already present at birth (Mehler et al., 1988; Morgan et al., 1987). Processing of auditory beat and meter has been demonstrated in premature newborns with a mean gestational age of 33 weeks (Edalati et al., 2023). The ability to perceive these cues is thought to facilitate the segmentation of speech input (Morgan et al., 1987) and to provide infants with an efficient mechanism for processing language across the lifespan (Gordon, Jacobs, et al., 2015).

A recent quantitative meta-analysis (Heard & Lee, 2020) provided evidence suggesting overlapping neural structures recruited in rhythm and grammar tasks, specifically in the brain areas involved in temporal hierarchy processing and potentially in predictive coding, or the act of generating predictions on what will occur (Friston, 2005). Such processes are likely implicated in rhythm and grammar processing: both domains entail the hierarchical organization of serial temporal input and the listener's active prediction of the next item to efficiently process upcoming events (Kuperberg & Jaeger, 2016; Persici et al., 2019; Vuust & Vitek, 2014). Moreover, research has shown that the ability to predict upcoming input is facilitated by rhythmic regularities (Andreou et al., 2011; Jones et al., 2002; Jones & Boltz, 1989; Large & Jones, 1999; Repp, 2005; Repp & Su, 2013). Crucially, both temporal hierarchy processing and predictive coding may be subserved by the neural tracking of rhythms (Arnal & Giraud, 2012; Fiveash et al., 2021; Ladányi, Persici, et al., 2020), a mechanism that we here term 'neural entrainment'.

Neural oscillations, produced by networks of neurons firing synchronously (Luo et al., 2006), are found to be entrained (i.e., synchronized) by the rhythms of higher-level structures such as meter (Fujioka et al., 2015) and syntactic phrases (Ding, Melloni, et al., 2017). This entrainment, which is already evident to beat and meter in 6- and 7-month-old infants (Cirelli et al., 2016; Flaten et al., 2022), is thought to be important for rhythm processing (Fujioka et al., 2015), but also for speech decoding, processing, and comprehension (Ahissar et al., 2001; Ding, Patel, et al., 2017; Ghitza, 2012; Luo & Poeppel, 2007), as well as for other processes including visual attention (see also Obleser & Kayser, 2019).

According to the Dynamic Attending Theory (DAT: Jones, 2019; Jones & Boltz, 1989; Large & Jones, 1999), auditory rhythms entrain attention rhythms, such that there are increases in attention at time points in the stimulus that are temporally regular and predictable (e.g., beats in music or stress in language). These attention rhythms may depend on the neural entrainment of multiple nested oscillations (Jones, 2019). Such nested entrainment has been observed in language processing. As shown by Ding, Melloni, Zhang, Tian, and Poeppel (2016), when English and Mandarin adult speakers hear a series of spoken sentences that contain elements that are presented at a fixed rate, peaks of spectral energy are produced at multiple levels: in correspondence to the word rate, to the phrasal unit rate, and to the sentence rate. This suggests that the brain not only responds to surface features of the acoustic waveform but also encodes syntactic units (see also Ding, Melloni, et al., 2017). These findings suggest commonalities in how the brain processes musical and linguistic structure.

1.2. Neural activity during rhythmic and syntactic processing in the beta and gamma frequencies

If neural entrainment subserves both rhythmic and syntactic processing (Ladányi, Persici, et al., 2020), individual differences in neural activity during rhythmic listening should also correspond to individual differences in hierarchical processing in language. To investigate neural responses to hierarchically-organized auditory rhythms, Iversen, Repp, and Patel (2009) examined adult participants' processing of rhythmic structures of two tones and one rest varying in accent placement—and thus, perceived beat structure—while they underwent magnetoencephalography (MEG). Experimental conditions in Iversen et al. (2009) included listening to two types of rhythmic sequences that varied in the location of intensity accents marking the beat (i.e., in whether the first or the second tone of the sequence were played louder), and two other conditions in which the participants were asked to imagine the same intensity accents. Both oscillatory responses in the beta and gamma frequency bands, and Event-Related Fields (ERF; the magnetic counterpart of event-related potentials; ERPs), were found to be enhanced at time points that corresponded to sounds that were heard (or imagined) as the beat, which suggests that brain responses are sensitive to beat interpretation.

Neural activity in the beta and gamma frequency appears also relevant for rhythmic processing in subsequent studies. Beta oscillations were found to be linked to beat anticipation and processing (Fujioka et al., 2009, 2012; Kasdan et al., 2020) and thought to index meter representation (Fujioka et al., 2015) and to reflect the coupling of brain activity originated in distant brain regions, such as the auditory and motor cortices (Bartolo et al., 2014). As beta activity rebounds prior to the next stimulus and peaks are found in correspondence to expected events (Fujioka et al., 2012), this frequency band is also believed to play a role in predictive processes (Arnal & Giraud, 2012; Leventhal et al., 2012), which are important for efficient rhythmic processing, given the periodicity and regularity of most rhythmic patterns (Kotz et al., 2018). Gamma band activity is also thought to be involved in rhythmic processing: peaks in evoked (phase-locked) gamma activity are found after the onset of expected tones, with power changes that are modulated in response to strong and weak beats (Zanto et al., 2005). Moreover, peaks in induced (non-phase-locked) gamma persist even when expected tones are omitted (Zanto et al., 2005). These findings have been interpreted as suggesting a role of gamma in the formation of temporal expectancies, in internal anticipatory timing, and thus in beat perception (Fujioka et al., 2009; Zanto et al., 2005; Zanto et al., 2005).

Both bands were also found to be implicated in grammar processing. Bastiaansen, Magyari, and Hagoort (2010) showed with MEG that a power increase in the beta frequency band occurs in correspondence to the timing of unification operations, that is, the process of combining individual word information into the representation of the whole utterance in adults. These results are consistent with the idea that beta is

also involved in the transmission of *content*-specific predictions (and thus, in predictive coding; Arnal & Giraud, 2012). Gamma has also been associated with language skills involving hierarchical processing in the study by Ding and colleagues (2016) mentioned above: the peaks in activity found at the syllable, phrasal, and sentential level suggest the existence of cortical networks supporting the concurrent encoding of multiple linguistic structures, including phrases and sentences.

1.3. The present study

In the present study we adapted Iversen et al. (2009)'s paradigm by presenting six-year-old children with rhythmic tone-tone-rest sequences differing in placement of an intensity accent marking the beat, while recording electroencephalography (EEG). The placement of an intensity accent either on the first or on the second tone of the sequence was expected to generate two distinct beat percepts. This expectation was based on prior results, using the same paradigm with typically-developed adults (Iversen et al., 2009) and adults with a neuro-developmental disorder associated with atypical sensory processing and attentional difficulties called Williams syndrome (Kasdan et al., 2020), and on research showing that neural responses to the beat are already detectable in infancy (Cirelli et al., 2016; Winkler et al., 2009).

As mentioned above, beat perception entails processing of temporal information in a hierarchical fashion and is thus an example of hierarchical processing in the music domain. In line with previous studies (e.g., Iversen et al., 2009; Kasdan et al., 2020), we assumed that greater efficiency in capturing differences in temporal patterns would reflect better beat processing and encoding. The same children were administered standardized tests of rhythm discrimination abilities and grammar skills. The aim was to investigate whether enhanced tracking of hierarchical structures in rhythm resulted in better manipulation of hiearchical structures in language. More specifically, we aimed to examine whether individual differences in neural responses to the beat were associated with and concurrently predicted individual differences in grammar task performance over and above behavioral rhythm discrimination performance (an effect previously shown in Gordon, Shivers, et al., 2015; Gordon, Jacobs, Schuele, & McAuley, 2015, as mentioned above). Children in this age range are an interesting population because they are still in the process of developing their grammar (Arndt & Schuele, 2013) and beat perception (Nave-Blodgett et al., 2021) abilities. Although they have reached peak cortical grey matter volume, they are still undergoing important changes in brain development (Bethlehem al., 2022). Investigating how individual differences in one domain relate to individual differences in the other domain in this population is particularly important for our understanding of the mechanisms underlying language development.

Children's neural activity was recorded while they listened passively to the auditory stimuli. Given the associations found between rhythm and language in other work (e.g., Chern et al., 2018; Gordon, Shivers, et al., 2015; Nitin et al., 2023) and the hypothesized shared mechanism for hierarchical processing between the two domains (Heard & Lee, 2020; Ladányi, Persici, et al., 2020), we expected individual differences in the magnitude of neural responses to these rhythms to relate to individual differences in grammar task performance, based on the assumption that greater neural responses represent more effective encoding of the rhythms. Neural responses were quantified by looking at each individual child's neural differentiation of different accent patterns, in beta and gamma bands, in order to arrive at EEG-based metrics of their effectiveness of differentiating beat patterns. In particular, we explored whether these evoked beta and gamma EEG patterns could explain variance in children's proficiency in two grammatical categories, Complex Syntax and Transformation over and above performance in a behavioral rhythm task. The use of EEG measures is theoretically and pragmatically justified. First, measuring rhythmic processing passively with EEG while participants watch a muted video should focus their attention on the video rather than the rhythm that

they listen to. Consequently, this indirect measure of rhythmic processing might give clear insight into rhythm discrimination than a task-driven behavioral approach. Secondly, if neural measures recorded passively were found to explain significant variance in grammar performance, these findings would motivate the use of such paradigms in infants and toddlers for the examination of their grammatical development and the early identification of language difficulties (Ladányi, Persici, et al., 2020).

The term 'Complex Syntax' refers to sentences that contain multiple clauses and thus have complex structural dependency relations (e.g., relative clauses: "the woman ate the fish *that her husband caught in the river*"). These utterances that include subordinate, infinite, complement, and relative clauses, are acquired between the ages of two and three years (Bloom et al., 1984). However, as opposed to simpler structures (Vasilyeva et al., 2008), growth in complex syntax continues throughout the school years (Arndt & Schuele, 2013). The term 'Transformation' (related to the syntactic operation called Movement; Radford, 1988) refers to uni-clausal sentences with a non-canonical word order (e.g., wh-interrogatives in English: "what did the woman eat?"), requiring the reordering of sentence elements before responding. Thus, both Transformation and Complex Syntax require individuals to perform complex grammatical operations. However, these operations are of greater complexity in the case of Complex Syntax, which comprises multi-clausal sentences with longer structural dependencies than Transformation (Diessel, 2004). See Fig. 1 for a visual depiction of these dependencies. The ability to perform such operations is fundamental for efficient language processing and development (Diessel, 2004). Rhythm discrimination abilities were tested in two tasks in which children were required to judge whether multiple successive presentations of a rhythm were the same or different. Note that these tasks tapping into children's general rhythm perception abilities rely on sequence learning and memory-based skills (Fiveash et al., 2022) and may potentially tap into a separable rhythm sequence-based construct from the beat perception task assessed with EEG.

Prior behavioral studies with school-aged children have shown that rhythmic processing skills relate to spoken grammar performance, both in general (Gordon, Shivers, et al., 2015) and in the two above-mentioned specific sub-categories (Gordon, Jacobs, et al., 2015; Lee et al., 2020). As discussed above, this link between rhythm perception and language production abilities may be due to shared biology via partially overlapping brain networks (Heard & Lee, 2020). According to recent theoretical accounts, behavioral (phenotypic) associations between musical and language traits may also partly result from genetic pleiotropy, or underlying common genetic architecture, among those traits (Fiveash et al., 2021; Ladányi, Persici, et al., 2020; Nayak et al., 2022). The possibility of uncovering overlapping underlying neurobiology across distinct skills/tasks is a particular motivation for the present work. Behaviorally, individual differences in hierarchical rhythmic ability may be associated with an advantage in spoken tasks. If hierarchical processing in both domains is subserved by neural entrainment, we hypothesize that individual differences in neural responses to musical rhythm (manifesting in peaks in beta and gamma at the onset of the beat) will correspond to individual differences in grammar ability, especially in the case of sentences that require more refined grammatical operations, i.e., in Complex Syntax.

2. Methods

2.1. Participants

This study reports EEG data of a cohort of participants for whom behavioral scores were previously reported (Gordon, Shivers, et al., 2015). Detailed participant characteristics and behavioral scores of the 25 children who participated (12 females, 13 males; 21 right-handed, 4 left-handed), ages between 5;11 and 7;1 years (mean: 6;6 years, *SD* = 4 months), are reported in Table 1. All data were collected on the same
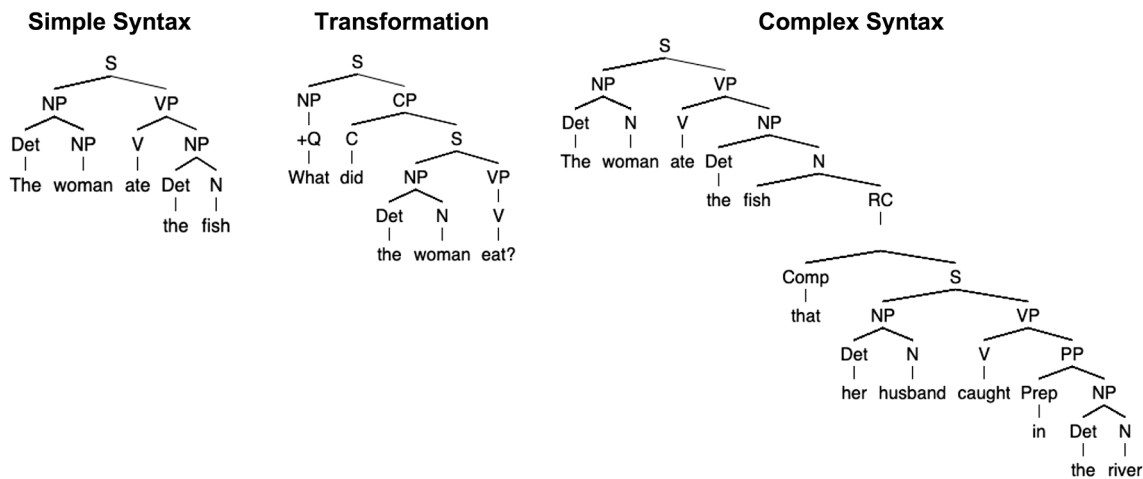
Fig. 1. Syntactic trees of sentence examples of simple syntax, transformation, and complex syntax.

Table 1

Participants' characteristics and results of the behavioral tests as reported in Gordon, Shivers et al., 2015.

	Mean (SD)	Percent correct (SD)
N	25	/
Age	6;6 (0;4)	/
Males: females	13:12	/
Maternal education level (9 levels)	7.32 (0.83)	/
Music experience (years)	1.08 (0.81)	/
Nonverbal intelligence (standard PTONI score)	110 (22.7)	/
SPELT-3 standard score	114 (6.87)	/
Transformation score (11 items)	9.52 (1.55)	86.54 (14.16)
Complex Syntax score (12 items)	9.6 (1.61)	81.33 (13.45)
BBA behavioral rhythm test (d'prime)	1.10 (0.76)	/
PMMA behavioral rhythm test	/	70.9 (12.6)

day for each child; participants were first tested in the EEG paradigm and then in the behavioral tests. All were native speakers of American English and, as reported by their parents, had normal hearing, normal neurological health, and typical language, cognitive, and emotional development. Mean maternal education (used as proxy for socio-economic status (SES)) corresponded to three or four years of under-graduate education; children had little music experience, on average corresponding to a year of music class in school. Non-verbal intelligence was measured using the Primary Test of Nonverbal Intelligence (PTONI; Ehrler & McGhee, 2008). See Table 1 for more information.

A sensitivity analysis was conducted in G*Power 3.1 (Faul et al., 2009) to examine the smallest effect size that our study was powered to detect with 80 % certainty, given our sample size and an alpha of 0.05. This test revealed that our study was powered to reliably detect an effect size of $f^2 = 0.35$, traditionally considered to be large (Cohen, 1988), in regression models with five to six predictors.

The study was approved by the Institutional Review Board of Van-derbilt University (Nashville, TN, United States of America), in agree-ment with the Declaration of Helsinki. Written informed consent and verbal assent were obtained from the parent and child respectively.

2.2. Behavioral assessments

2.2.1. Expressive grammar skills

Grammar ability was tested with the Structured Photographic Expressive Language Test (SPELT-3; Dawson et al., 2003), in which children are presented with various photographs and asked questions designed to elicit specific syntactic constructions such as relative or interrogative clauses. For example, in one test item, children are shown a picture and prompted with the sentence, 'The lady has some cookies. What is she asking the children?'; children are expected to answer, 'Do you want some cookies?'. In another test item, children are shown a picture of two girls and prompted with the sentence 'One sister went to school early and one left home late. Who did not arrive on time?', and are expected to answer: 'The girl who was late/who left home late/that was late'. Appropriate responses thus require both efficient compre-hension of the verbal probe delivered by the experimenter and devel-oped expressive abilities. Standard (age-normed) total scores and grammatical category mean scores for Complex Syntax and Trans-formation items are reported in Gordon, Jacobs, et al. (2015) and in Table 1. Standard (age-normed) total scores showed a performance within the normal range for all participants (mean: 114.16; *SD* = 6.87, range: 98–125, indicating typical development).

2.2.2. Musical rhythm skills

Rhythm discrimination abilities were tested with two computer-based games: the children's beat-based advantage assessment (BBA; originally adapted from adult work on beat perception by Grahn & Brett, 2009), and the rhythm section of the Primary Measures of Music Audi-ation (PMMA; Gordon, 1979). In both tests children were asked to judge whether multiple successive rhythm presentations were the same or different. BBA includes 28 test trials, including both simple and complex rhythms; PMMA includes 40 test trials. Both tasks were used because of their unique attributes: only PMMA is a standardized test of music aptitude; on the other hand, BBA addresses beat perception specifically (e.g., Fiveash et al., 2022, Grahn and Brett, 2009, and Niarchou et al., 2022, for previous papers using this task), whereas PMMA includes stimuli with more varied metrical structures (see Gordon, Shivers, et al., 2015 for more details on the two tasks). A behavioral Rhythm Composite score was calculated by averaging z-transformed d́ scores on the BBA assessment (commonly used in signal detection studies; Macmillan & Creelman, 2005) and z-transformed PMMA percent correct scores (note that the BBA and PMMA z-transformed scores were moderately and significantly correlated; *r* = 0.43, *p* =.030). The Rhythm Composite measure previously reported in Gordon, Shivers, et al. (2015) was ob-tained by averaging percent correct scores from both tests. Here we decided to use d́ scores for the BBA assessment, because these scores are unaffected by response bias and thus convey a more precise measure of participants' beat perception skills compared to percent correct scores. The new variable was highly correlated with the previously reported Rhythm Composite measure (*r* = 0.99, *p* <.001). See Gordon, Shivers,

et al. (2015) for more details on the assessments.

2.3. Electroencephalography

2.3.1. Stimuli

The auditory sequences in the EEG paradigm were adapted from the physical accent conditions used in Iversen, Repp, and Patel (2009); see Fig. 2. This paradigm was chosen because it has proven successful in yielding neural responses to the beat in the beta and gamma frequency ranges in previous research on adults (Iversen et al., 2009; Kasdan et al., 2020). Here the aim was to investigate whether the same paradigm could be applied to children. In this study we did not include the imagined beat conditions (for which top-down control is important) used in Iversen et al.'s study for theoretical and practical purposes. Most importantly, we expected each of physical accent condition patterns to elicit different beat percepts (i.e., perception of patterns with different beat positions). This would be the case if children's early latency neural responses to rhythms are affected by physical stimuli, a phenomenon that has been previously shown using the same paradigm with typically developed adults (Iversen et al., 2009) and with adults with Williams syndrome (Kasdan et al., 2020). We thus predicted that fluctuations in early ERP sensory responses and in beta and gamma evoked activity would reflect differentiation between the two rhythmic sequences. Second, a shorter experiment with only implicit listening allowed keeping the session at a duration that reduced variance in children's ability to follow task demands. In each of the two physical accent conditions, tones had a frequency of 1000 Hz, a duration of 50 ms, an intensity of 58 dB, and an inter-onset interval of 200 ms between the first and second tone in each pair (the tone-tone-rest pattern thus had a duration of 600 ms). In the Accent1 condition, sequences had a strong–weak-rest pattern: the accent, conveyed by increasing the intensity of the sound by 10 dB, occurred on the first tone; in the Accent2 condition, the accent occurred on the second tone, thus creating a weak-strong-rest pattern (see Fig. 2). Each block consisted of 50 repetitions of the same pattern for a duration of 30 s and was presented nine times per condition in random order (for an overall task duration of nine minutes). In the analyses, we excluded the first two trials in each block and analyzed only the following forty-eight repetitions, for a total of 864 trials for analysis.

2.4. Procedure

2.4.1. EEG acquisition

Participants' EEG recordings were collected individually in a sound-dampened room at the EEG Lab at the Vanderbilt Kennedy Center in Nashville, Tennessee (USA). Brain activity was recorded continuously using 128 Ag/AgCl electrodes embedded in soft sponges (EGI Geodesic Sensor Net, Eugene, OR, USA). Lower eye channels were only available on some sizes of the EEG nets and consequently were excluded from further analyses. EEG signals were sampled at 500 Hz for temporal precision (the data of two participants were inadvertently sampled at 250 Hz and then later upsampled to 500 Hz). Data was acquired in Net Station 4.4 with high-impedance amplifier NetAmps 200 and with a Butterworth hardware filter of 0.1 to 200 Hz; impedances were adjusted

to below 40 kΩ before the start of the paradigm. Stimuli were presented at 72 dB through a single speaker placed above the participant's head. EEG was recorded while participants listened passively to the auditory stimuli; no behavioral responses were required. To keep participants engaged but still during data acquisition, an age-appropriate video with muted sound was shown. Participants were asked to watch the video and stay still and that they would "hear some sounds"; no further instructions were given. The entire session lasted about 45 min and included an additional experiment for a separate project.

2.4.2. EEG data preprocessing

Data processing was performed using the EEGLAB toolbox (Delorme & Makeig, 2004) in MATLAB R2017b (The Mathworks Inc, 2017). Signals were smoothed using a 100-Hz low-pass filter and a 0.5-Hz high-pass filter to eliminate non-brain-related frequencies. Line noise was cleaned using the pop_cleanline function. Bad electrodes characterized by consistently high noise levels were identified using the Artifact Subspace Reconstruction (ASR) approach and interpolated using the spherical spline interpolation algorithm (Perrin et al., 1989). Data were re-referenced to the average of all channels.

Artifacts in the data were identified using Independent Component Analysis (ICA); the type of artifact was identified with the help of ICLabel (Pion-Tonachini et al., 2019). Cardiac and ocular artifacts were manually removed (range: 0–3 components (heart, eye blink or sideways eye movement) per participant; mean = 1.48). Next, data were divided into 1600-ms epochs with the interval of [-400, +1200], time-locked to the onset of the first tone (tone-tone-rest, in both conditions); a wide window is needed for time–frequency analyses to avoid window-edge artifacts (Roach & Mathalon, 2008). Epochs were rejected if they exceeded a $-100/+100$ μV threshold, as in Kasdan et al. (2020). An average of 7.38 % ($SD = 10.25$ %) of epochs were rejected. A technical error caused a subset of $n = 13$ participants to have received an extra two blocks of trials during data acquisition; these were retained during data clean-up to improve ICA reliability (Debener et al., 2010) and then discarded in the following way. First, we computed the number of epochs kept after cleaning of the $n = 12$ that received the correct number of trials and found that it was 407 trials for Accent1 and 406 trials for Accent2. To prevent the other $n = 13$ subset of data from having a disproportionate impact on the signal-to-noise ratio, we removed the last 75 clean epochs from 12 of those 13 participants for further analysis (the remaining 13th participant had only 291 and 220 trials remaining after artifact rejection respectively and thus no further trials were removed). In the final dataset of $N = 25$, the mean number of epochs was not different between conditions ($t(24) = 0.605$, $p = .551$, $SD = 19.84$; for Accent1: mean trials = 409 ($SD = 33$) and Accent2: mean trials = 407 ($SD = 46$)). Note that analysis including extra epochs, for the purpose of comparison, did not significantly affect any of the final results; p-values and cluster latencies were virtually identical (see Table 5S in the Supplementary materials).

2.4.3. Data analysis

Similar to prior work using this paradigm (Iversen et al., 2009; Kasdan et al., 2020), we conducted ERP and time–frequency analysis on the EEG data using the FieldTrip Toolbox (Oostenveld et al., 2011).

Accent1 Condition
Physical accent on 1st tone

Accent2 Condition
Physical accent on 2nd tone

Fig. 2. Stimulus conditions in the EEG paradigm (used with permission from Kasdan et al., 2020).

Statistical analyses on brain-behavior relationships were carried out using R (R Core Team, 2022).

2.4.3.1. Event-related potentials (ERPs). Event-related potential (ERPs) were generated from the average across trials for each condition and utilized for evoked time–frequency analyses (see below). ERPs were then trimmed to [-100 to 700 ms], baseline-corrected with a baseline of [-100, 0] ms and low-pass filtered with a Butterworth filter of 30 Hz for subsequent ERP analyses. Line noise was further removed with the DFT filter.

2.4.3.2. Time-frequency representations (TFRs). Evoked (phase-locked) activity for each condition was obtained by convolving the average ERP waveform with a family of Morlet wavelets with a width of six cycles, from 12 Hz to 50 Hz, with a frequency step of 1 Hz and a time step of 2 ms in the time window [-200, +800], with zero being the onset of each tone-tone-rest pattern (see Tallon-Baudry et al., 1996, and Herrmann et al., 2005, for more details on this analysis). Note that this wide window was only used for TFR calculation; overlapping latencies were not tested statistically (see below). Power values for the resulting time frequency representations (TFRs) were then normalized (baseline-corrected) at each time point, frequency, and channel as relative power change in relation to the mean total power averaged across time points and across both conditions separately for each frequency and channel, to compensate for inter-individual variability in absolute power.

2.4.3.3. Cluster-based permutation statistical analyses. Statistical analyses were conducted in FieldTrip, using a cluster randomization procedure to identify consistent trends in activity in clusters of channels and using cluster-based permutation tests coded for within-subjects design to test for differences in activity between conditions (Maris & Oostenveld, 2007). For ERPs, the Accent1 vs. Accent2 conditions were compared in the time window [-100, +500] (to avoid testing overlapping latencies) at each time point and channel using a dependent-samples t-test. All the data points from adjacent channels with significant t-values at $p < .025$ (two-tailed) were used to generate clusters; all the other data points were zeroed. Adjacency was determined via the distance between neighboring channels using the FieldTrip triangulation parameter, and a minimum of two neighboring channels at a given time point and channel were required for that data point to be included in a cluster (i.e., three channels were the minimum to meet the adjacency criteria). The cluster-level test statistic for each cluster was then calculated by summing all the t-values within that cluster. The significance of each cluster was assessed using a Monte Carlo method: randomly permuted values from both conditions were taken and pooled together to form two new subsets (simulated dummy conditions), which were used to obtain a dummy cluster test statistic. The same process was repeated 5,000 times, to obtain 5,000 dummy clusters. The resulting cluster p-values are the proportions of permutations in which the true cluster statistic exceeds the simulated cluster (dummy) statistic. Clusters with a cluster $p < .05$ were considered statistically significant. The same procedure was used for time–frequency representations (TFRs) of Accent1 vs. Accent2, for average power within two frequency bands: 13–23 Hz for beta, 24–50 Hz for gamma.

2.4.3.4. Individual differences and brain-behavior correlations. Variables representing the difference of neural activity between conditions were obtained for TFRs in each frequency band, using a method previously reported by Lense et al. (2014): power values (for time–frequency data) of the difference between conditions at each time point and channel in the latency band (for clusters that surpassed group-level statistical significance) are summed for each participant. In other words, the group-level cluster parameters for the spatial and temporal distribution of the cluster were applied to the individual data at each participant and the sum of the difference of values at each time and point and condition

are computed. Thus, for each participant this procedure resulted in a single value for each of the four cluster-sums (i.e., for the beta and gamma clusters for the Beat1 and Beat2 effects). Potential differences in the magnitude of activity between cluster-sums within the same frequency band were investigated using paired t-tests.

To determine if neural responses to rhythm were associated with individual differences in spoken grammar skill, we performed Spearman correlations between significant beta and gamma evoked clusters and the total SPELT-3 score or the scores in each of the two SPELT-3 sub-categories of interest (all z-scored), controlling for age, non-verbal intelligence, music experience, and maternal education (as a proxy for SES) and with p-values adjusted for multiple comparisons (Benjamini-Hochberg correction). Note that the Beat1 effect clusters correspond to increased neural responses around the onset of the first tone (for Accent1 versus Accent2), and the Beat2 effect clusters correspond to larger neural responses around the onset of the second tone (for Accent2 versus Accent1). We focused on beta and gamma TFRs in relation to language task performance, given prior literature on the role of beta and gamma in rhythmic attending (see Introduction). The correlations between beta and gamma activity and between Beat1 and Beat2 effect clusters are reported in Table 3S in the Supplementary materials.

Hierarchical regressions were then used to assess the potential contribution of neural markers of beat processing to grammar over and above the variance explained by behavioral musical rhythm (which was previously reported in Gordon, Shivers, et al., 2015; Gordon, Jacobs, Schuele, & McAuley, 2015). Models were compared using F-tests (Field, Miles, & Field, 2012). The variance explained by each model was examined by looking at Adjusted R^2 values. The analysis of individual differences in ERP amplitude (calculated following the same procedure as for TFRs) and of their relationship with syntactic scores was carried out as exploratory work and is reported in the Supplementary materials.

3. Results

3.1. Cluster-based analysis of ERPs

Results of cluster-based permutation tests revealed two significant clusters in correspondence to Accent1 vs. Accent2 at the onset of the first tone ("Beat1 effect"; see Fig. 3) and three significant clusters in correspondence to Accent2 versus Accent1 at the onset of the second tone ("Beat2 effect"; see Fig. 3). Detailed ERP cluster results are reported in Table 2. Specifically, we found a negative ERP deflection in the neural activity over frontal regions of the scalp when the accent was on the first tone (ERP-Beat1effect-early-negativity-frontal: from −46 to 54 ms, cluster $p = .009$; Fig. 3A), and a posterior positive effect around the same time window (ERP-Beat1effect-early-positivity-posterior; from −30 to 50 ms, cluster $p = .024$; Fig. 3B). Given that the clusters occur at the same latencies and display nearly identical waveforms with opposite polarities, they might be two extremes of the same dipole. We also found three significant clusters showing a Beat2 effect: a frontal negative effect (ERP-Beat2effect-late-negativity-frontal1 from 222 to 462 ms, cluster $p < .001$; Fig. 3C), a posterior positive effect (ERP-Beat2effect-late-positivity-posterior from 214 and 410 ms, $p < .001$; Fig. 3D), and, finally, a negative frontal effect (ERP-Beat2effect-late-negativity-frontal2) in the time window between 424 and 500 ms, cluster $p = .032$ (Fig. 3E). Following the same reasoning stated above, we interpret the first two Beat2 effect clusters as the two extremes of the same dipole, and thus likely reflecting the same cortical source. The similar topographies of the two frontal Beat2 effect clusters also suggest that they might be the same brain source. Taken together, the early latencies of four out of five of the ERP clusters (with respect to beat onset) suggest that early sensory responses are modulated by beat perception.

3.2. Cluster-based analysis of TFRs

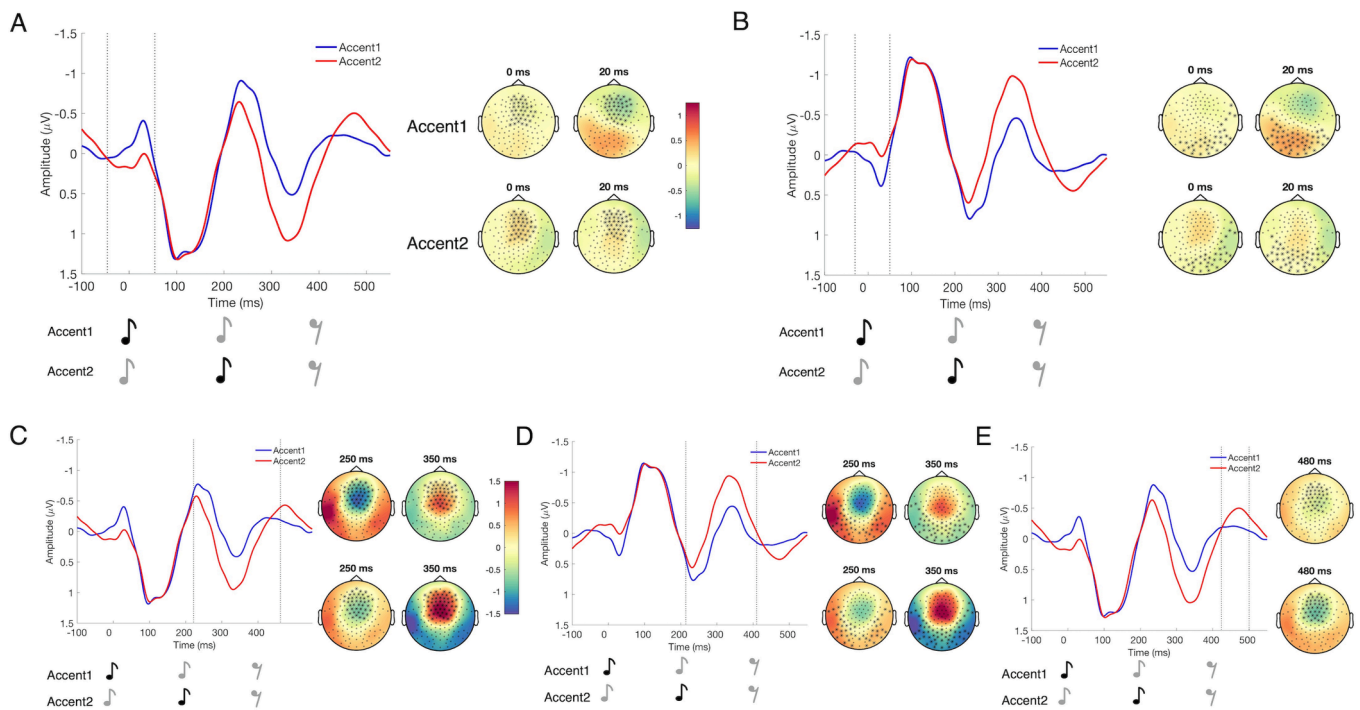Results of cluster-based permutation tests revealed significant

Fig. 3. Grand average ERP (N = 25) Beat 1 (A, B) and Beat 2 (C, D, E) effects. The dotted lines indicate the start and end latencies of the significant clusters found in correspondence to the first beat (Beat1 effect; onset of the sequence at 0 ms) or in correspondence to the second beat (Beat2 effect; onset of the sequence at 0 ms). ERPs for condition Accent1 (strong–weak-rest) are denoted in blue, and ERPs for condition Accent2 (weak-strong-rest) are in red. ERP plots (on the left of each panel) show amplitude in microvolts averaged over all the channels belonging to the significant cluster (y-axis) with negative up by convention, and time in ms (x-axis). Topographic plots (on the right) show scalp distribution of amplitude at time points indicated, which were chosen to correspond to peaks within the cluster window. Channels belonging to the significant cluster are indicated with asterisks. Panel A shows the ERP-Beat1effect-early-negativity-frontal cluster over fronto-central regions (cluster p =.009). Panel B shows ERP-Beat2effect-early-positivity-posterior cluster over posterior central regions (cluster p =.024). Panel C shows the ERP-Beat2effect-late-negativity-frontal1 cluster over fronto-central regions (cluster p <.001). Panel D shows the ERP-Beat2effect-late-positivity-posterior cluster over posterior central regions (cluster p <.001). Panel E shows the ERP-Beat2effect-late-negativity-frontal2 cluster over fronto-central regions (cluster p =.032). (For interpretation of the references to color in this figure legend, the reader is referred to the web version of this article.)

Table 2
Characteristics of the significant clusters identified by cluster-based permutation tests of Event-Related Potentials (ERP).

Cluster label	Latency	Scalp distribution	Cluster p-value
ERP-Beat1effect-early-negativity-frontal	−46 to 54 ms	Frontal	0.009
ERP-Beat1effect-early-positivity-posterior	−30 to 50 ms	Posterior	0.024
ERP-Beat2effect-late-negativity-frontal1	222 to 462 ms	Frontal	< 0.001
ERP-Beat2effect-late-positivity-posterior	214 to 410 ms	Posterior	< 0.001
ERP-Beat2effect-late-negativity-frontal2	424 to 500 ms	Frontal	0.032

evoked beta and gamma activity in correspondence to the beat onset in each condition, which was greater for the accented beat. Specifically, significantly greater neural activity in beta band with a widespread scalp distribution (EEG-beta-Beat1effect cluster p <.001; Fig. 4A) and in gamma band with a posterior-central distribution (EEG-gamma-Beat1effect cluster p =.033; Fig. 4B) were found in correspondence to the first tone of the Accent1 versus Accent2 condition (Beat1 effect). Similarly, greater activity in the beta band (cluster p =.006; Fig. 5A) and in gamma band (cluster p =.028; Fig. 5B), with both clusters distributed over fronto-central regions, were found in correspondence to the second tone in the Accent2 condition versus Accent1 condition (Beat2 effect; see Fig. 5). Cluster latencies are reported in Table 3.

Paired t-tests showed that beta activity was significantly larger in the

Beat1 effect cluster than in the Beat2 effect cluster ($t(24) = 3.975$, p <.001). Gamma activity was not significantly different between clusters (p >.05).

3.3. Brain-behavior relationships

3.3.1. Neural markers of beat processing and expressive language ability: Correlations

The analysis of beta and gamma activity yielded significant associations with language scores, as reported on Table 3 and as shown in Fig. 6. For the beta band, the results showed strong and significant correlations between the EEG-beta-Beat1effect cluster and SPELT-3 total (Fig. 6A), and also with both the Transformation (Fig. 6B) and Complex Syntax (Fig. 6C) sub-scores.

In the gamma band, a strong and significant correlation was also found between the EEG-gamma-Beat1effect cluster and Complex Syntax (Fig. 6D). No significant correlations were found between the EEG-beta-Beat2effect cluster and SPELT-3 (whether total or in each subcategory; SPELT-3 total: p =.838; Transformation: p =.838; Complex Syntax: p =.759), between the EEG-gamma-Beat1effect cluster and SPELT-3 total (p =.107) or Transformation (p =.195), or between the EEG-gamma-Beat2effect cluster and any of the SPELT-3 categories (SPELT-3 total: p =.107; Transformation: p =.303) or Complex Syntax (p =.324).

To summarize, a larger Beat1 effect (suggesting enhanced neural responses to the first tone when it was accented) was positively associated with language scores, in alignment with our hypothesis. Yet, a larger Beat2 effect (suggesting enhanced neural responses to the second tone when it was accented) was not significantly associated with language scores, contrary to our hypothesis. We then turned to examine
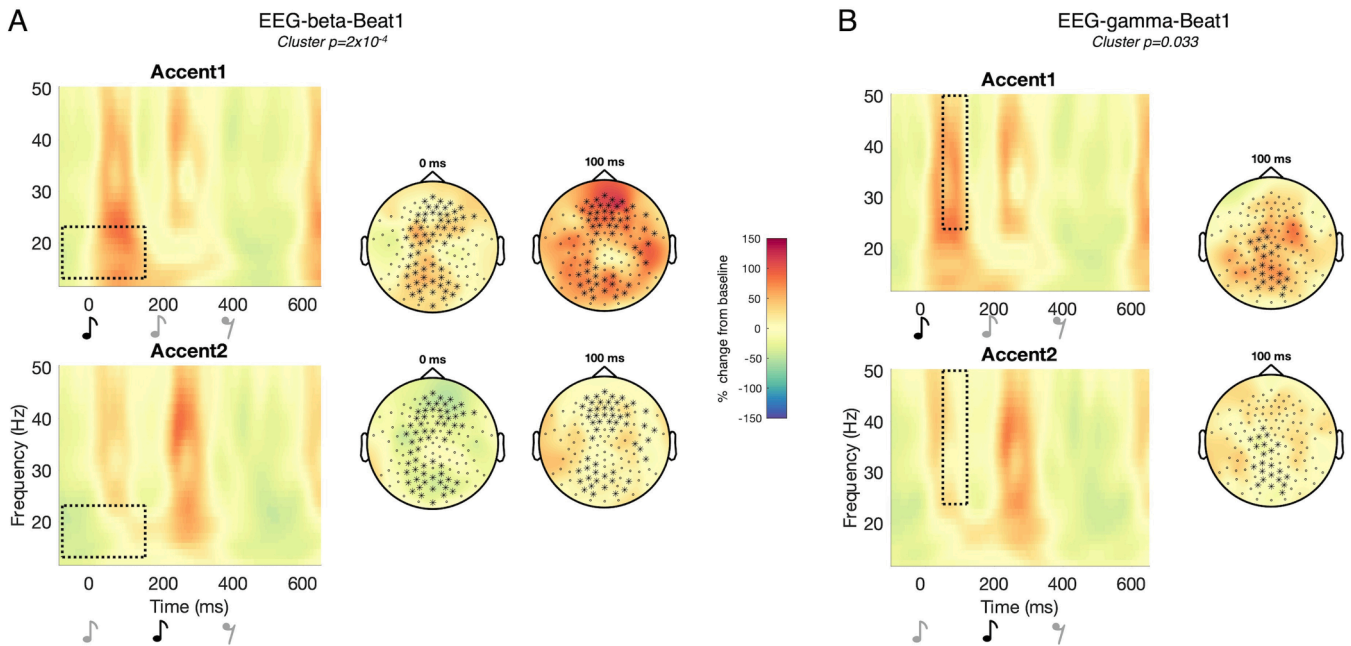
Fig. 4. Beat 1 effect. Time-frequency representations (TFRs) and topographies for grand average (N = 25), showing normalized power changes in EEG evoked beta (panel A) and gamma (panel B) neural activity relative to the first tone (Beat1 effect). TFRs are shown on the left side of each panel and are the average of all the channels belonging to the cluster; the black dashed box indicates the time and frequency boundaries of the significant clusters. Representative topographies (on the right side of each panel) are shown at the indicated latencies within the clusters. Color scale represents the percent change from baseline (i.e., normalized power). Significant channels are marked with black asterisks. As signaled by the darker red colors, increased beta and gamma neural activity is found at the first tone when this was accented (in the Accent1 condition (top row) as compared to the Accent2 condition (bottom row)). (For interpretation of the references to color in this figure legend, the reader is referred to the web version of this article.)
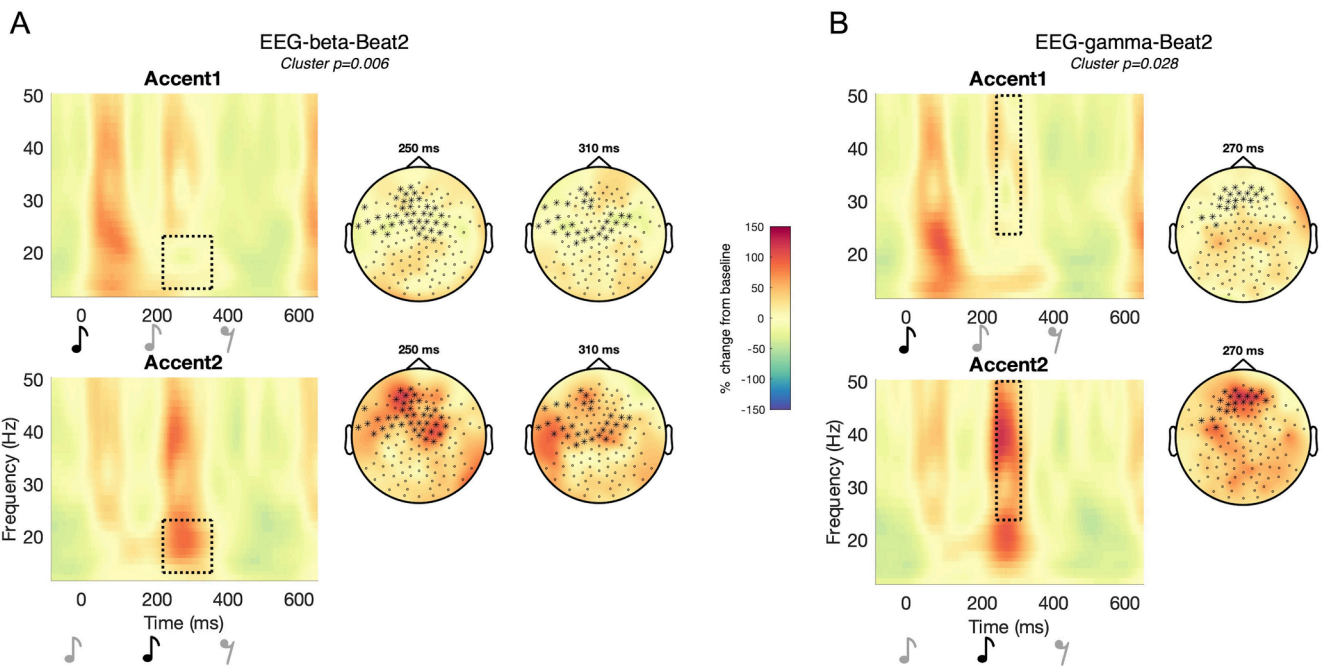

Fig. 5. Beat 2 effect. TFRs and topographies for EEG evoked beta (panel A) and gamma (panel B) neural activity relative to the second tone (Beat2 effect). See Fig. 4 for description of representation. Increased beta and gamma neural activity is found for the second tone when this was accented (in the Accent2 condition (bottom row) as compared to the Accent1 condition (top row)).

effects of brain variables on grammar scores over and above variances accounted for by rhythm behavioral variables.

3.3.2. Hierarchical regressions

As reported in Gordon, Shivers, et al. (2015) and Gordon, Jacobs, Schuele, and McAuley (2015), the Rhythm Composite score (calculated here as the mean of the z-scored BBA d' and PMMA percent correct) explained unique variance in SPELT-3 scores, whether total or within the Transformation and Complex Syntax subcategories. Here, we added the significant beta and gamma evoked clusters to see if additional

Table 3

Characteristics of the significant clusters identified by cluster-based permutation tests of beta and gamma evoked activity during the EEG rhythm task, and their significant associations with individual differences in language skills.

Cluster label	Latency	Scalp distribution	Cluster p-value	Significant association with language skills		
				Grammar category	$r_{(s)}$	p
EEG-beta-Beat1 effect	−92 to 146 ms	Widely distributed	< 0.001	SPELT-3 total	0.60	0.018
				Transformation	0.56	0.030
				Complex Syntax	0.63	0.014
EEG-beta-Beat2 effect	214 to 352 ms	Fronto-central	0.006	–		
EEG-gamma-Beat1 effect	50 to 120 ms	Posterior-central	0.033	Complex Syntax	0.62	0.014
EEG-gamma-Beat2 effect	238 to 306 ms	Fronto-central	0.028			

Note. P-values were adjusted for multiple comparisons using the Benjamini-Hochberg correction.
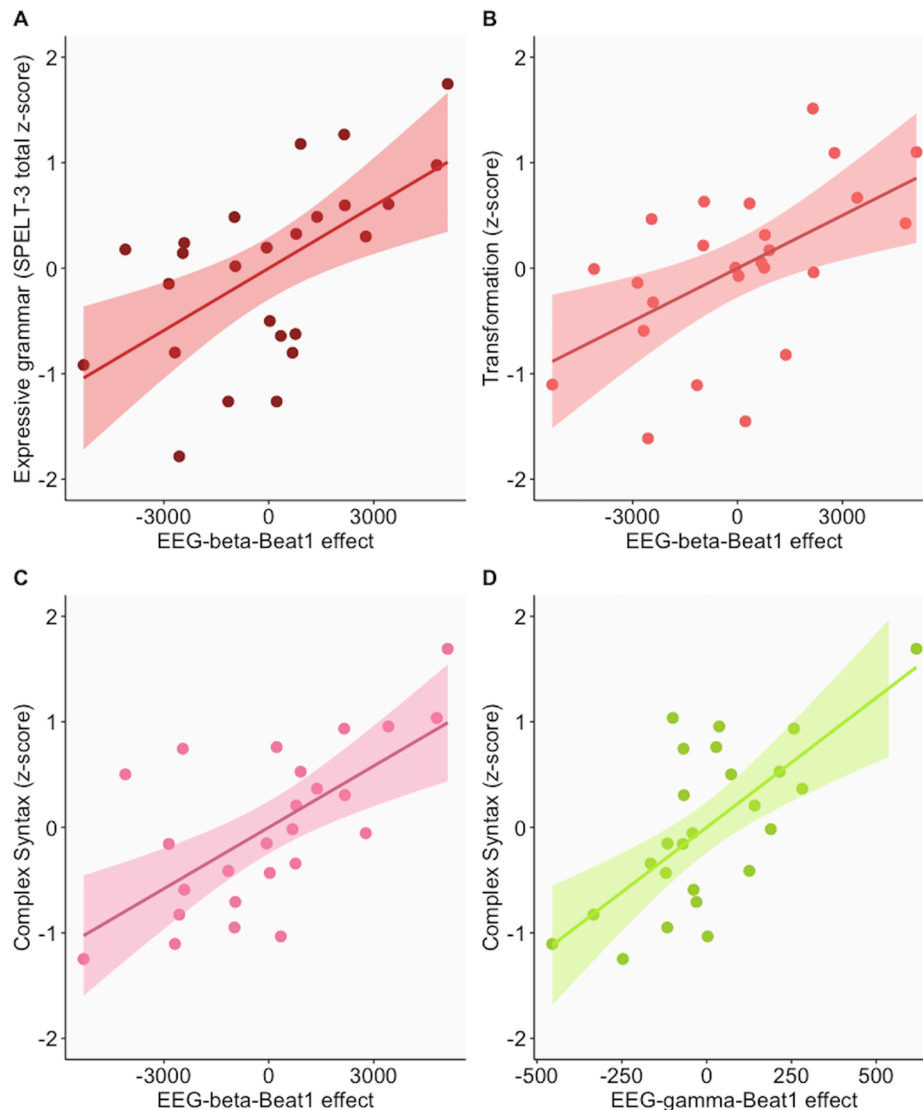

Fig. 6. Statistically significant partial correlations between Z-transformed expressive grammar scores on the SPELT-3 and evoked variables, controlling for age, nonverbal IQ, music experience, and maternal education. All partial correlations are graphed as the correlation between the residuals of two linear regression models with the covariates as predictors and the expressive grammar score or evoked cluster of interest as dependent variable. Panels A, B, and C show the significant partial correlations between the EEG-beta-Beat1effect and SPELT-3, Transformation, and Complex Syntax, respectively. Panel D shows the significant partial correlation between EEG-gamma-Beat1effect and Complex Syntax. Note that the Beat1effect corresponds to larger responses near the onset of the first tone in Accent1. Thus, higher, positive numbers indicate larger neural activity. See Methods and Results for more detail. *p <.05, **p <.01, ***p <.001.

variance in the grammar scores of our participants could be explained by their neural responses to rhythm over and above that explained by rhythm behavior.

Table 4 shows the fit and the results of the base models and of the models in which a significant effect of our neural measures was found.

Table 4
Model fit and summary of the hierarchical regression base models and of the models showing a significant effect of the neural variables.

Dependent Variable	Model	Model Fit						Predictor	Standardized beta	S.E.	t-value	p-value
		F	*df*	*p*	Adjusted R^2	AIC	BIC					
SPELT-3 total score	Model 1 (base model)	4.799	5, 19	0.005	0.442	63.509	72.041	**Rhythm Composite Score**	**0.658**	**0.180**	**3.664**	**0.002**
	Model 2	10.719	6, 18	< 0.001	0.708	47.923	57.674	**Rhythm Composite Score**	**0.615**	**0.130**	**4.723**	**<0.001**
								EEG-beta-Beat1effect	**0.512**	**0.119**	**4.286**	**<0.001**
	Model 4	7.946	6, 18	< 0.001	0.635	53.568	63.319	**Rhythm Composite Score**	**0.566**	**0.148**	**3.827**	**0.001**
								EEG-gamma-Beat1effect	**0.468**	**0.141**	**3.320**	**0.004**
Transformation score	Model 1 (base model)	3.940	5, 19	0.013	0.380	66.140	74.672	**Rhythm Composite Score**	**0.417**	**0.189**	**2.204**	**0.040**
	Model 2	6.308	6, 18	0.001	0.570	57.619	67.370	**Rhythm Composite Score**	**0.380**	**0.158**	**2.403**	**0.027**
								EEG-beta-Beat1effect	**0.445**	**0.145**	**3.069**	**0.007**
Complex Syntax score	Model 1 (base model)	3.135	5, 19	0.031	0.308	68.887	77.419	Rhythm Composite Score	0.276	0.200	1.378	0.184
	Model 2	6.769	6, 18	0.001	0.591	56.411	66.162	Rhythm Composite Score	0.231	0.154	1.495	0.152
								EEG-beta-Beat1effect	**0.532**	**0.142**	**3.757**	**0.001**
	Model 4	6.990	6, 18	0.001	0.600	55.851	65.602	Rhythm Composite Score	0.164	0.155	1.059	0.304
								EEG-gamma-Beat1effect	**0.569**	**0.148**	**3.853**	**0.001**

Note. Each base model included age, nonverbal IQ, music experience, maternal education, and the Rhythm Composite score. Model 2, 3, 4, and 5 additionally included one of our neural variables (EEG-beta-Beat1effect, EEG-beta-Beat2effect, EEG-gamma-Beat1effect, EEG-gamma-Beat2effect). Significant effects are in bold. Adjusted R^2 indicates the proportion of variance explained adjusted for the number of predictors in the model. AIC = Akaike Information Criterion. BIC = Bayesian Information Criterion.

The full results of these analyses[1] are reported in Table 4S in the Supplementary materials. Adding the EEG-beta-Beat1effect cluster to the base model (i.e., Model 1, which included age, non-verbal intelligence, music experience, maternal education, and the behavioral rhythm score as predictors) increased significantly the ability of the model to fit the data ($F(1, 18) = 18.373$, $p <.001$; see Table 4 for AIC and BIC values). The more complex model (i.e., Model 2) explained 26.6 % more variance in the total SPELT-3 score than Model 1 (the Adjusted R^2 value changed from 0.442 in Model 1 to 0.708 in Model 2). Although the EEG-beta-Beat1effect cluster ($\beta = 0.512$) had less impact on the total SPELT-3 scores than the Rhythm Composite score ($\beta = 0.615$) in Model 2, its addition still contributed to significantly increase the amount of variance explained in the grammatical scores. Adding the EEG-gamma-Beat1effect cluster to the base model also improved significantly model fit ($F(1, 18) = 11.021$, $p =.004$). The variance explained in the total SPELT-3 scores this time was larger by 19.3 % (the Adjusted R^2 value changed from 0.442 in Model 1 to 0.635 in Model 4). Again, in the final model, the Rhythm Composite score was the most influential predictor ($\beta = 0.566$; β of the EEG-gamma-Beat1effect cluster = 0.468). On the contrary, the EEG-beta-Beat2effect and EEG-gamma-Beat2effect clusters did not significantly contribute to explain unique variance in these scores (beta: $p =.409$, gamma: $p =.376$). See Fig. 7 to see the standardized beta coefficients and their 95 % confidence intervals in these models.

Similar results were found for the Complex Syntax scores as for total SPELT-3 scores. Complex Syntax results showed significant ($p =.001$) and unique variance (28.3 %; the Adjusted R^2 value changed from 0.308 in Model 1 to 0.591 in Model 2) explained by the EEG-beta-Beat1effect

cluster ($F(1, 18) = 14.118$, $p =.001$). In this model, the added neural variable was also the most influential ($\beta = 0.532$; Rhythm Composite score $\beta = 0.231$). Adding the EEG-gamma-Beat1effect cluster to the base model also increased significantly ($F(1, 18) = 14.846$, $p =.001$) the proportion of variance explained, this time by 29.2 % (Adjusted R^2 value change from 0.308 in Model 1 to 0.600 in Model 4). Again, the added neural variable was the most influential predictor of the children's Complex Syntax scores ($\beta = 0.569$; Rhythm Composite score $\beta = 0.164$). No significant results were found when adding the EEG-beta-Beat2effect ($p =.243$) or EEG-gamma-Beat2effect ($p =.533$) clusters.

For the Transformation scores, results showed that the EEG-beta-Beat1effect cluster explained significant ($p =.007$) and 19 % unique variance (the R^2 value changed from 0.380 in Model 1 to 0.570 in Model 2; $F(1, 18) = 9.419$, $p =.007$). The effect of the EEG-beta-Beat1effect cluster ($\beta = 0.445$) on the Transformation scores was stronger than that of the Rhythm Composite score ($\beta = 0.380$). However, none of the other evoked variables significantly contributed to explain unique variance in these scores (the EEG-beta-Beat2effect: $p =.189$; EEG-gamma-Beat1effect: $p =.094$; EEG-gamma-Beat2effect: $p =.821$).

4. Discussion

The aim of the present study was to identify neural oscillatory markers of beat processing that may relate to individual differences in language performance in children. We hypothesized that the strength of neural responses to the beat would be linked to individual differences in language skills, especially in hierarchical processing, based on previous behavioral findings indicating that musical rhythm discrimination skills relate to expressive and receptive grammar performance (Gordon, Shivers, et al., 2015; Gordon, Jacobs, Schuele, & McAuley, 2015; Swaminathan & Schellenberg, 2019). Neural responses to the beat point to processes of neural entrainment, which are crucial for hierarchical processing in both rhythm and language (Ahissar et al., 2001; Ding,
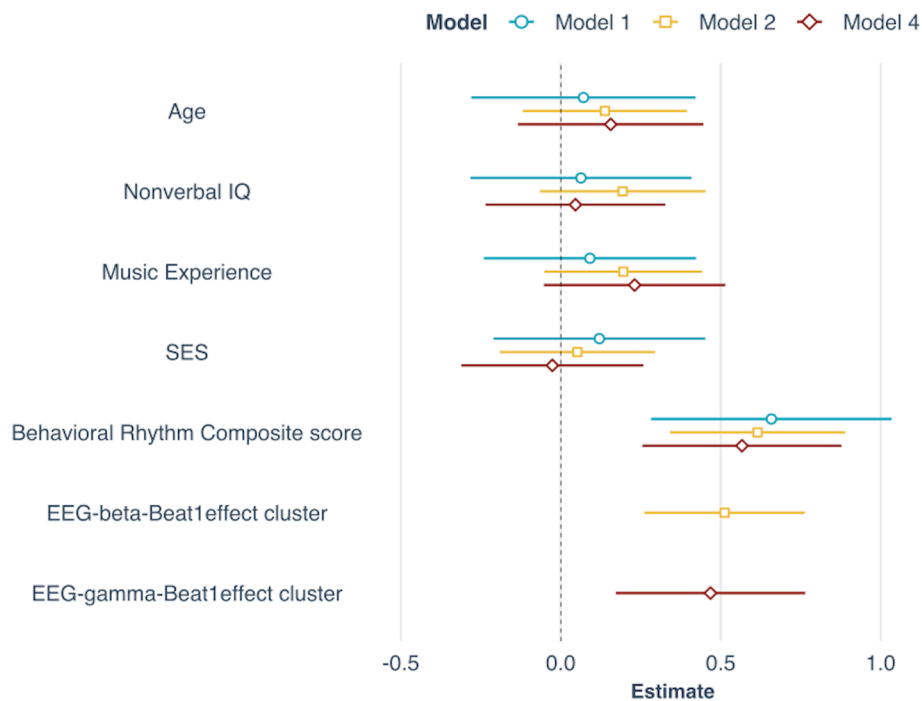
---

[1] Note that running these analyses using the scores obtained in the validated, standardized rhythm discrimination test (i.e., PMMA) as predictor, rather than Rhythm Composite scores, does not significantly affect the results relative to the effects of the evoked clusters.

Fig. 7. Standardized beta coefficients and 95% confidence intervals from the models on SPELT-3 scores. Model 1 includes age, nonverbal IQ, music experience, maternal education (SES), and the behavioral Rhythm Composite score as predictors. Models 2 and 4 additionally include the EEG-beta-Beat1effect and EEG-gamma-Beat1effect clusters, respectively.

Patel, et al., 2017; Ghitza, 2012; Ladányi, Persici, et al., 2020; Luo & Poeppel, 2007); moreover, individual differences in neural entrainment may bootstrap acquisition of hierarchically organized structure during language development. Infants and newborns can process and neurally track rhythmic regularities in musical stimuli (Cirelli et al., 2016; Flaten et al., 2022) as well as in speech (Kalashnikova et al., 2019). Optimal entrainment, both internally between multiple oscillators (Jones, 2019) and to rhythm in music and grammatical structure in language, may help scaffold hierarchical processing by facilitating temporal integration of smaller units into larger linguistic structures (Ding et al., 2016). Neural mechanisms of entrainment may thus account for a relative strength in a given child's ability to process hierarchical structures in music (Large & Jones, 1999) and to handle complex sentence structures during language acquisition.

Our study showed that the neural responses of six-year-old children peaked in correspondence to rhythmic auditory stimuli, converging with previous research findings in adults showing modulations of evoked beta, gamma, and ERF neural responses to different physically-accented beat patterns (e.g., Iversen et al., 2009); taken together, this reinforces the idea that beat perception is robust from an early age (if not innate; see Winkler et al., 2009). Specifically, the children showed neural responses that fluctuated in both ERPs (with components that are typical of auditory stimulation; Ponton et al., 2000) and evoked beta and gamma activity according to the beat pattern that they heard: enhanced neural responses were found at the strong beat (i.e., the moment that the tone received the physical accent). Moreover, beyond the contribution of behaviorally measured rhythm perception skills, our measures of the differentiation of the children's neural responses in the beta and gamma bands between accented and unaccented beats in initial position (termed Beat1 effect) were found to correlate and explain variance in spoken syntactic abilities. This was especially the case for task items requiring more refined complex syntactic operations (i.e., generation of multi-clausal sentences with dependency relations), for which the most influential predictors were our beta and gamma neural measures. The hierarchical regression model also included age, nonverbal IQ, music experience, and SES variables, ruling out the possibility that these

results are due only to general maturity and cognitive effects or to environmental influences.

Our study is the first to our knowledge to show that individual differences in expressive syntactic performance are explained by neural – beyond and complementarily to behavioral – measures of rhythm in children. The unique variance explained by the beta and gamma neural measures recorded around the onset of the first tone, which lies between 19 and 29.2 %, significantly contributes to the ability of the model to predict the children's concurrent syntactic performance scores, for a total of variance explained that reaches 70.8 % in one case (see Model 2 on SPELT-3 total scores in Table 4). Importantly, this indicates that individual differences in grammar skills may be concurrently predicted by passively collected neural measures; the fact that beat perception is already measurable in newborns (Winkler et al., 2009) suggests that such paradigms may also be used to investigate grammatical development in children from a very early age. Of note is that spoken syntactic performance was related to beta activity (in correspondence to accents on the first tone), while gamma activity appeared to be of less impact. We consider that these results may be linked to the fact that beat-related effects have been found to be stronger in the beta band as defined here (Fujioka et al., 2009, 2012; Iversen et al., 2009; Snyder & Large, 2005) and to the nature of the language task that measured children's expressive language skills: beta activity has been shown to be strongly linked to motor processes (Salmelin et al., 1995); relations between neural activity and language performance would be interesting to explore in the context of comprehension skills (Ladányi et al., 2023). Moreover, larger effects of beta than gamma frequency corresponding to the first tone might result from the fact that we only found a first-tone advantage for the former one. The mean beta activity was significantly larger in the Beat1 effect cluster than in the Beat2 effect cluster, and no significant differences were found in the gamma frequency range. As each cluster-sum variable was obtained by subtracting activity to the unaccented tone from activity to the beat, these results suggest increased beta activity in response to the beat when it is in first vs. second position. These findings suggest a first-tone advantage in the beta but not gamma frequency range that are consistent with prior results (Iversen et al.,

2009).

Contrary to our expectations, larger neural activity around the onset of the second tone seemed not to reflect better syntactic encoding: activity in this latency did not significantly contribute to explain variance in the children's grammar scores; moreover, the gamma results showed that the measures of the differentiation of the children's neural responses between accented and unaccented beats in second position (termed Beat2 effect) were not significantly associated with their syntactic scores. In other words, this means that those children who better discriminated the difference between accented and unaccented tones in second position did not necessarily exhibit stronger grammar abilities. These results, together with those showing the children's facilitation in relation to the first tone, could potentially be related to the most common stress patterns found in English. The strong–weak (or trochaic) pattern in speech (Cutler & Carter, 1987) is preferred by English speakers already at infancy (Echols et al., 1997; Jusczyk et al., 1993) and is acquired earlier than weak-strong (iambic) patterns (Ballard et al., 2012). Children's ability to discriminate accent information on the second tone may thus be less relevant to linguistic structure processing in English. Another potential explanation is linked to surface-level characteristics of the stimuli and/or serial-order effects, which might have rendered Beat2 effects smaller, and the association with language not to emerge statistically. As mentioned above, beta responses to beats in second position appear to be more reduced than those to beats in first position. Reduced beta responses to the second tone in a sequence may be due to the shorter refractory time before the onset of the second tone (IOI: 200 ms) than before the onset of the first tone (IOI: 400 ms) and/or to the suppression of the neural response to the repetition of the same stimulus (Grill-Spector et al., 2006). Future research investigating other languages and manipulating metrical structures and IOIs should clarify the source of these findings.

The significant brain-behavior correlations observed in relation to the evoked time–frequency neural activity indicate that investigating high-frequency phase-locked oscillatory activity provides new insight into neural markers that could help explain behavioral associations between rhythm processing and grammar performance reported elsewhere (Gordon, Shivers, et al., 2015; Gordon, Jacobs, Schuele, & McAuley, 2015; Lee et al., 2020; Politimou et al., 2019; Swaminathan & Schellenberg, 2019). In the context of meter, these findings align with prior findings of beta (which is thought to be closely related to rhythmic processing (Fujioka et al., 2012) and to beat perception (Cirelli et al., 2016; Flaten et al., 2022; Fujioka et al., 2015), as well as with prior findings of gamma activity (which has been shown to track hierarchical linguistic structures: Ding et al., 2016). It should be noted that the beta band is also thought to play an important role in coordinating the entrainment of the auditory and motor cortices (Bartolo et al., 2014), which are important for both language and rhythm. Several neurophysiological processes therein may contribute to these brain-behavioral associations, including sensorimotor coupling in auditory-motor networks and fine-grained auditory processing (Fiveash et al., 2021; Ladányi, Persici, et al., 2020). Optimal entrainment between the auditory and motor cortices may facilitate rhythm and language processing, possibly by enhancing predictions about sensory events (Kotz et al., 2009, Kotz & Schmidt-Kassow, 2015).

Beta and gamma neural activity have indeed both been associated with predictive processes regarding the timing (when) and content (what) of upcoming events (Arnal & Giraud, 2012; Leventhal et al., 2012; Zanto et al., 2005), as they are thought to reflect synchronization of neuronal populations during sensory processing of regular, predictive information (Arnal & Giraud, 2012), which is possibly further shaped by top-down, learned modulation from the motor system (Iversen et al., 2009; Patel & Iversen, 2014). Beyond sensory processing of the physical accents, it appears in the current study that more efficient synchronization and beat structure differentiation in the beta and gamma bands is associated with extraction and parsing of hierarchically complex structures (Ding et al., 2016). These new EEG results, and the rapid time

course at which they unfold in tracking the onset of a musical beat, complement recent fMRI meta-analysis findings of overlapping neural resources of complex syntax and musical rhythm in adulthood (Heard & Lee, 2020). The findings of the current study are particularly relevant in consideration of the ontogeny of complex syntax use, which has a notably protracted developmental acquisition throughout school-age in children with typical development (Vasilyeva et al., 2008) and becomes increasingly vital for the communication of complex academic material, life skills, and social relationships (Brimo et al., 2017; Fujiki et al., 1999). The link between rhythm and language via neural oscillations is also reinforced by the finding that musicians – who show more precise endogenous neural entrainment to rhythm (Stupacher et al., 2017) – find it easier to acquire new hierarchical structures in an artificial language (Brod & Opitz, 2012). The association between hierarchical structure extraction and parsing in rhythm and language fit well in the context of evolutionary arguments as to the role of rhythm in human communication and interaction (Honing, 2018; Patel & Iversen, 2014; Savage et al., 2020). Humans appear to be naturally disposed to perceiving, producing, and appreciating rhythms across a variety of contexts (Savage et al., 2015), including verbal and nonverbal social interaction (Cirelli et al., 2018). In children, the predisposition to perceive, extract, and manipulate hierarchical information may guide language learning by facilitating scaffolding of incoming input into distinct language units (Flaten et al., 2022) and thus ease processing and manipulation of complex linguistic information.

5. Conclusions

To conclude, the present study showed how individual differences in neural markers of beat perception relate to individual differences in expressive grammar in six-year-old children. Individual differences in mastery of complex syntax structure were especially related to beta band activity. In line with previous research (e.g., Politimou et al., 2019; Swaminathan & Schellenberg, 2019; Woodruff Carr et al., 2014), the present findings suggest that precise neural encoding of temporally organized rhythmic structure may reflect a mechanism biologically related to typical language development and acquisition. Impairment of mechanisms linked to neural markers of rhythm may eventually have critical implications for bolstering earlier identification and treatment of developmental speech and language disorders (see the Atypical Rhythm Risk Hypothesis by Ladányi, Persici, et al., 2020).

It is important to note that future investigations in a larger sample size and utilizing an in-depth, comprehensive battery of complex syntax language assessments, investigating both comprehension and production, are warranted to assess the generalizability of these results. As indicated by a sensitivity analysis, the current study was not powered to reliably detect small and medium-sized effects. This suggests that more subtle effects may have not emerged due to its small sample size. Furthermore, the same analysis suggests that the effect size of the relation of beta and gamma activity in response to the first tone on the SPELT-3 total scores and on transformation scores, respectively, are only of medium size and should therefore be interpreted with caution.

In addition, it could be argued that that the beat patterns in our EEG task were fairly simple compared to the hierarchical structures included in our grammar task; future work may include beat patterns with more complex hierarchical structures to investigate whether individual differences in complex beat tasks explain even more variance in children's grammar abilities. Incorporating imagined beat conditions in the future may also enlarge our understanding of language processes: in the present study we only examined passive (and possibly pre-attentive) evoked responses, but studies have suggested that attentive and top-down processes also contribute to the discovery of temporal structure and provide potential additional vistas that might reveal novel insights into the development of the neural underpinnings of the temporal processing of language. The inclusion of such conditions would also help clarify whether the neural responses observed in this study are partly

influenced by the physical properties of the stimuli. Although previous work suggest that these types of responses are neural correlates of beat tracking (Iversen et al., 2009; Kasdan et al., 2020), possible effects due to perception of intensity cannot be excluded and direct claims about predictive processing cannot be made. In addition, including conditions in which inter-onset-intervals are jittered to disrupt beat perception or are equal before the onset of the two tones, and/or conditions with metrical structures of varying complexity in future work may help disentangle the interplay between sensory and cognitive processing and establish whether the relationship between neural activity and grammar production skills observed here is driven by beat perception specifically.

Future studies could also explore whether the same relationship holds if more traditional paradigms tapping into pre-attentive processes (such as oddball paradigms) are used to investigate children's neural responses to the beat. Research with other languages may also be of help in investigating whether these results are language-specific or generalizable across different languages. Importantly, the current study shows an association between beat and grammar processing using concurrent measures; to clarify whether this relationship is causal, it will be important to further probe if actively supporting rhythm development (for instance, through participation in music programs) leads to transfer effects in grammar development in longitudinal designs. Finally, the use of similar paradigms in adults may provide insight into whether hierarchical processing is a mechanism underlying both rhythm and grammar processing and is not only driving language learning in development. While the same language test employed in the current study may not be sufficiently complex to highlight individual differences in adults, tasks requiring more sophisticated syntactic operations may reveal similar relations to those observed here. These expectations are based on prior findings of associations between activity in the beta (Bastiaansen et al., 2010) and gamma (Ding et al., 2016) frequency range and extraction and processing of syntactic structures, between rhythm and grammar processing in several tasks (Canette et al., 2019; Kotz et al., 2009; Schmidt-Kassow & Kotz, 2008), and between artificial grammar learning and musical expertise (Brod & Opitz, 2012).

Nonetheless, the current study is one of the few ones so far to apply an individual differences approach toward understanding of children's neural activity during a musical task, and furthermore to show, in a developmental population, evidence of a link between individual differences in neural activity and in a trait in a different domain. These promising findings suggest that this approach to rhythmic neural processing with EEG in children may not only be a passive, implicit, rapid, and relatively inexpensive way to assay a particular dimension of musical beat processing that appears to have strong links with spoken language development. Similar approaches may also serve as an important tool for predicting individual differences in complex language use in school-age children with a wide range of language skills.

Declaration of Competing Interest

Data availability

The link to de-identified EEG and behavioral data and to data analysis scripts is in the paper.

Acknowledgements

Author contributions

Conceptualization of Study: RLG, JDM, SAK. Study Design: RLG, JDM, SAK, APK, VP, SDB. Materials and Methods: RLG, JDM, JRI, SAK. Data Analysis: VP, RLG, SDB. Interpretation of Primary Findings: VP, RLG, JDM. Writing, Editing, and Reviewing Drafts: VP, RLG, JDM, SAK, APK, JRI, SDB.

Data Accessibility Statement

De-identified EEG and behavioral variables and data analysis scripts for brain-behavioral correlations and hierarchical regressions are available at https://doi.org/10.17605/OSF.IO/CSFK2.

Appendix A. Supplementary material

Supplementary data to this article can be found online at https://doi.org/10.1016/j.bandl.2023.105345.

References

Ahissar, E., Nagarajan, S., Ahissar, M., Protopapas, A., Mahncke, H., & Merzenich, M. M. (2001). Speech comprehension is correlated with temporal response patterns recorded from auditory cortex. *Proceedings of the National Academy of Sciences of the United States of America, 98*(23), 13367–13372. https://doi.org/10.1073/pnas.201400998

Andreou, L. V., Kashino, M., & Chait, M. (2011). The role of temporal regularity in auditory segregation. *Hearing Research, 280*(1–2), 228–235. https://doi.org/10.1016/j.heares.2011.06.001

Arndt, K. B., & Schuele, C. M. (2013). Multiclausal utterances aren't just for big kids. A framework for analysis of complex syntax production in spoken language preschool- and early school-age children. *Topics in Language Disorders, 33*(2), 125–139. https://doi.org/10.1097/TLD.0b013e31828f9ee8

Arnal, L. H., & Giraud, A.-L. (2012). Cortical oscillations and sensory predictions. *Trends in Cognitive Sciences, 16*(7), 390–398. https://doi.org/10.1016/j.tics.2012.05.003

Ballard, K. J., Djaja, D., Arciuli, J., James, D. G. H., & van Doorn, J. (2012). Developmental trajectory for production of prosody: Lexical stress contrastivity in children ages 3 to 7 years and in adults. *Journal of Speech, Language, and Hearing Research, 55*(6), 1822–1835. https://doi.org/10.1044/1092-4388(2012/11-0257)

Bartolo, R., Prado, L., & Merchant, H. (2014). Information processing in the primate basal ganglia during sensory-guided and internally driven rhythmic tapping. *The Journal of Neuroscience, 34*(11), 3910–3923. https://doi.org/10.1523/JNEUROSCI.2679-13.2014

Bastiaansen, M., Magyari, L., & Hagoort, P. (2010). Syntactic unification operations are reflected in oscillatory dynamics during on-line sentence comprehension. *Journal of Cognitive Neuroscience, 22*(7), 1333–1347. https://doi.org/10.1162/jocn.2009.21283

Beier, E. J., & Ferreira, F. (2018). The temporal prediction of stress in speech and its relation to musical beat perception. *Frontiers in Psychology, 9*, 431. https://doi.org/10.3389/fpsyg.2018.00431

Bethlehem, R. A. I., Seidlitz, J., White, S. R., et al. (2022). Brain charts for the human lifespan. *Nature, 604*, 525–533. https://doi.org/10.1038/s41586-022-04554-y

Bloom, L., Tackeff, J., & Lahey, M. (1984). Learning *to* in complement constructions. *Journal of Child Language, 11*, 391–406. https://doi.org/10.1017/S0305000900005833

Brimo, D., Apel, K., & Fountain, T. (2017). Examining the contributions of syntactic awareness and syntactic knowledge to reading comprehension. *Journal of Research in Reading, 40*(1), 57–74. https://doi.org/10.1111/1467-9817.12050

Brod, G., & Opitz, B. (2012). Does it really matter? Separating the effects of musical training on syntax acquisition. *Frontiers in Psychology, 3*(DEC). https://doi.org/10.3389/fpsyg.2012.00543

Canette, L.-H., Bedoin, N., Lalitte, P., Bigand, E., & Tillmann, B. (2019). The regularity of rhythmic primes influences syntax processing in adults. *Auditory Perception & Cognition, 2*(3), 163–179. https://doi.org/10.1080/25742442.2020.1752080

Chern, A., Tillmann, B., Vaughan, C., & Gordon, R. L. (2018). New evidence of a rhythmic priming effect that enhances grammaticality judgments in children. *Journal of Experimental Child Psychology, 173*, 371–379. https://doi.org/10.1016/j.jecp.2018.04.007

Cirelli, L. K., Einarson, K. M., & Trainor, L. J. (2014). Interpersonal synchrony increases prosocial behavior in infants. *Developmental Science, 17*(6), 1003–1011. https://doi.org/10.1111/desc.12193

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*(MAY), 1–11. https://doi.org/10.3389/fnins.2016.00229

Cirelli, L. K., Trehub, S. E., & Trainor, L. J. (2018). Rhythm and melody as social signals for infants. *Annals of the New York Academy of Sciences, 1423*, 66–72. https://doi.org/10.1111/nyas.13580

Cohen, J. (1988). *Statistical power analysis for the behavioral sciences*. Hillsdale, NJ: Erlbaum.

Cutler, A., & Carter, D. (1987). The predominance of strong initial syllables in the English vocabulary. *Computer Speech and Language, 2*, 133–142. https://doi.org/10.1016/0885-2308(87)90004-0

Dawson, J., Stout, C., & Eyer, J. (2003). *Structured Photographic Expressive Language Test-3*. Dekalb, IL: Janelle Publications.

Debener, S., Thorne, J., Schneider, T. R., & Campos Viola, F. (2010). Using ICA for the Analysis of Multi-Channel EEG Data. In M. Ullsperger & S. Debener (Eds.), Simultaneous EEG and fMRI: recording, analysis and application (pp. 121–133). Oxford Univerity Press.

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

Diessel, H. (2004). *The Acquisition of Complex Sentences. The Acquisition of Complex Sentences*. Cambridge University Press. Doi: 10.1017/cbo9780511486531.

Dilley, L. C., & McAuley, J. D. (2008). Distal prosodic context affects word segmentation and lexical processing. *Journal of Memory and Language, 59*(3), 294–311. https://doi.org/10.1016/j.jml.2008.06.006

Ding, N., Melloni, L., Yang, A., Wang, Y., Zhang, W., & Poeppel, D. (2017). Characterizing neural entrainment to hierarchical linguistic units using electroencephalography (EEG). *Frontiers in Human Neuroscience, 11*, 481. https://doi.org/10.3389/fnhum.2017.00481

Ding, N., Melloni, L., Zhang, H., Tian, X., & Poeppel, D. (2016). Cortical tracking of hierarchical linguistic structures in connected speech. *Nature Neuroscience, 19*(1), 158–164. https://doi.org/10.1038/nn.4186

Ding, N., Patel, A. D., Chen, L., Butler, H., Luo, C., & Poeppel, D. (2017). Temporal modulations in speech and music. *Neuroscience and Biobehavioral Reviews, 81*, 181–187. https://doi.org/10.1016/j.neubiorev.2017.02.011

Echols, C. H., Crowhurst, M. J., & Childers, J. B. (1997). The perception of rhythmic units in speech by infants and adults. *Journal of Memory and Language, 36*(2), 202–225. https://doi.org/10.1006/jmla.1996.2483

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. *Journal of Neuroscience*, JN-RM-1100-22. Advance online publication. Doi: 10.1523/JNEUROSCI.1100-22.2023.

Ehrler, D. J., & McGhee, R. L. (2008). *PTONI: Primary Test of Nonverbal Intelligence*. Austin, TX: Pro-Ed.

Faul, F., Erdfelder, E., Buchner, A., & Lang, A.-G. (2009). Statistical power analyses using G*Power 3.1: Tests for correlation and regression analyses. *Behavior Research Methods, 41*, 1149–1160. https://doi.org/10.3758/BRM.41.4.1149

Field, A., Miles, J., & Field, Z. (2012). *Discovering Statistics Using R*. London, UK: SAGE Publications Ltd.

Fitch, W. T. (2013). Rhythmic cognition in humans and animals: Distinguishing meter and pulse perception. *Frontiers in Systems Neuroscience, 7*, 68. https://doi.org/10.3389/fnsys.2013.00068

Fitch, W. T. (2017). Empirical approaches to the study of language evolution. *Psychonomic Bulletin and Review, 24*(1), 3–33. https://doi.org/10.3758/s13423-017-1236-5

Fitch, W. T., & Martins, M. D. (2014). Hierarchical processing in music, language, and action: Lashley revisited. *Annals of the New York Academy of Sciences, 1316*(1), 87–104. https://doi.org/10.1111/nyas.12406

Fiveash, A., Bedoin, N., Gordon, R. L., & Tillmann, B. (2021). Processing rhythm in speech and music: Shared mechanisms and implications for developmental speech and language disorders. *Neuropsychology, 35*(8), 771–791. https://doi.org/10.1037/neu0000766

Fiveash, A., Dalla Bella, S., Bigand, E., Gordon, R. L., & Tillmann, B. (2022). You got rhythm, or more: The multidimensionality of rhythmic abilities. *Attention, Perception, & Psychophysics, 84*, 1370–1392. https://doi.org/10.3758/s13414-022-02487-2

Flaten, E., Marshall, S., Dittrich, A., & Trainor, L. (2022). Evidence for top-down metre perception in infancy as shown by primed neural responses to an ambiguous rhythm. *European Journal of Neuroscience, 55*(8), 2003–2023. https://doi.org/10.1111/ejn.15671

Friston, K. (2005). A theory of cortical responses. *Philosophical Transactions of the Royal Society, B: Biological Sciences, 360*(1456), 815–836. https://doi.org/10.1098/rstb.2005.1622

Fujiki, M., Brinton, B., Hart, C. H., & Fitzgerald, A. H. (1999). Peer acceptance and friendship in children with Specific Language Impairment. *Topics in Language Disorders, 19*(2), 34–38.

Fujioka, T., Ross, B., & Trainor, L. J. (2015). Beta-Band Oscillations Represent Auditory Beat and Its Metrical Hierarchy in Perception and Imagery. *The Journal of Neuroscience, 11*(NOV), 15187–15198. https://doi.org/10.3389/fnins.2016.00389

Fujioka, T., Trainor, L. J., Large, E. W., & Ross, B. (2009). Beta and gamma rhythms in human auditory cortex during musical beat processing. *Annals of the New York Academy of Sciences, 1169*, 89–92. https://doi.org/10.1111/j.1749-6632.2009.04779.x

Fujioka, T., Trainor, L. J., Large, E. W., & Ross, B. (2012). Internalized timing of isochronous sounds is represented in neuromagnetic β oscillations. *The Journal of Neuroscience: The Official Journal of the Society for Neuroscience, 32*(5), 1791–1802. https://doi.org/10.1523/JNEUROSCI.4107-11.2012

Ghitza, O. (2012). On the role of theta-driven syllabic parsing in decoding speech: Intelligibility of speech with a manipulated modulation spectrum. *Frontiers in Psychology, 3*, 238. https://doi.org/10.3389/fpsyg.2012.00238

Gordon, E. E. (1979). *Primary Measures of Music Audiation*. Chicago, IL: GIA Publications.

Gordon, R. L., Jacobs, M. S., Schuele, C. M., & McAuley, J. D. (2015). Perspectives on the rhythm-grammar link and its implications for typical and atypical language development. *Annals of the New York Academy of Sciences, 1337*(1), 16–25. https://doi.org/10.1111/nyas.12683

Gordon, R. L., Shivers, C. M., Wieland, E. A., Kotz, S. A., Yoder, P. J., & Devin McAuley, J. (2015). Musical rhythm discrimination explains individual differences in grammar skills in children. *Developmental Science, 18*(4), 635–644. https://doi.org/10.1111/desc.12230

Grahn, J. A., & Brett, M. (2009). Impairment of beat-based rhythm discrimination in Parkinson's disease. *Cortex, 45*, 54–61. https://doi.org/10.1016/j.cortex.2008.01.005

Grill-Spector, K., Henson, R., & Martin, A. (2006). Repetition and the brain: Neural models of stimulus-specific effects. *Trends in cognitive sciences, 10*(1), 14–23. https://doi.org/10.1016/j.tics.2005.11.006

Heard, M., & Lee, Y. S. (2020). Shared neural resources of rhythm and syntax: An ALE Meta-Analysis. *Neuropsychologia, 137*, Article 107284. https://doi.org/10.1016/j.neuropsychologia.2019.107284

Herrmann, C. S., Grigutsch, M., & Busch, N. A. (2005). EEG oscillations and wavelet analysis. In T. C. Handy (Ed.), *Event-Related Potentials: A Methods Handbook* (pp. 229–259). Cambridge, MA: MIT Press.

Honing, H. (2018). On the biologica basis of musicality. *Annals of the New York Academy of Sciences, 1423*(1), 51–56. https://doi.org/10.1111/nyas.13638

Iversen, J. R., Repp, B. H., & Patel, A. D. (2009). Top-down control of rhythm perception modulates early auditory responses. *Annals of the New York Academy of Sciences, 1169*, 58–73. https://doi.org/10.1111/j.1749-6632.2009.04579.x

Jones, M. R. (2019). *Time will tell: A theory of Dynamic Attending*. New York: Oxford University Press.

Jones, M. R., & Boltz, M. (1989). Dynamic Attending and Responses to Time. *Psychological Review, 96*(3), 459–491. https://doi.org/10.1037/0033-295X.96.3.459

Jones, M. R., Moynihan, H., MacKenzie, N., & Puente, J. (2002). Temporal aspects of stimulus-driven attending in dynamic arrays. *Psychological Science, 13*(4), 313–319. https://doi.org/10.1111/1467-9280.00458

Jusczyk, P. W., Cutler, A., & Redanz, N. J. (1993). Infants' Preference for the Predominant Stress Patterns of English Words. *Child Development, 64*(3), 675–687. https://doi.org/10.1111/j.1467-8624.1993.tb02935.x

Kalashnikova, M., Goswami, U., & Burnham, D. (2019). Sensitivity to amplitude envelope rise time in infancy and vocabulary development at 3 years: A significant relationship. *Developmental Science, 22*, e12836.

Kasdan, A., Gordon, R. L., & Lense, M. D. (2020). Neurophysiological correlates of dynamic beat tracking in individuals with Williams syndrome. *Biological Psychiatry: Cognitive Neuroscience and Neuroimaging, 7*(11), 1183–1191. https://doi.org/10.1016/j.bpsc.2020.10.003

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

Kotz, S. A., & Schmidt-Kassow, M. (2015). Basal ganglia contribution to rule expectancy and temporal predictability in speech. *Cortex; a journal devoted to the study of the nervous system and behavior, 68*, 48–60. https://doi.org/10.1016/j.cortex.2015.02.021

Kotz, S. A., & Schwartze, M. (2010). Cortical speech processing unplugged: A timely subcortico-cortical framework. *Trends in Cognitive Sciences, 14*(9), 392–399. https://doi.org/10.1016/j.tics.2010.06.005

Kotz, S. A., & Schwartze, M. (2016). Motor-timing and sequencing in speech production: A general-purpose framework. In G. Hickok, & S. L. Small (Eds.), *Neurobiology of Language* (pp. 717–724). Academic Press. https://doi.org/10.1016/B978-0-12-407794-2.00057-2.

Kotz, S. A., Schwartze, M., & Schmidt-Kassow, M. (2009). Non-motor basal ganglia functions: A review and proposal for a model of sensory predictability in auditory language perception. *Cortex, 45*(8), 982–990. https://doi.org/10.1016/j.cortex.2009.02.010

Kuperberg, G. R., & Jaeger, T. F. (2016). What do we mean by prediction in language comprehension? *Language, Cognition, and Neuroscience, 31*, 32–59. https://doi.org/10.1080/23273798.2015.1102299

Ladányi, E., Lukács, Á., & Gervain, J. (2021). Does rhythmic priming improve grammatical processing in Hungarian-soeaking children with and without developmental language disorder? *Developmental Science, 24*(6), e13112.

Ladányi, E., Novakovic, M., Boorom, O. A., Aaron, A. S., Scartozzi, A. C., Gustavson, D. E., et al. (2023). Using motor tempi to understand rhythm and grammatical skills in developmental language disorder and typical language development. *Neurobiology of Language, 4*(1), 1–28. https://doi.org/10.1162/nol_a_00082

Ladányi, E., Persici, V., Fiveash, A., Tillmann, B., & Gordon, R. L. (2020). Is atypical rhythm a risk factor for developmental speech and language disorders? *Wiley Interdisciplinary Reviews: Cognitive Science, 11*(5), e1528.

Lang, M., Bahna, V., Shaver, J. H., Reddish, P., & Xygalatas, D. (2017). Sync to link: Endorphin-mediated synchrony effects on cooperation. *Biological Psychology, 127*, 191–197. https://doi.org/10.1016/j.biopsycho.2017.06.001

Large, E. W., & Jones, M. R. (1999). The dynamics of attending: How people track time-varying events. *Psychological Review, 106*(1), 119–159. https://doi.org/10.1037/0033-295X.106.1.119

Lashley, K. S. (1951). The problem of serial order in behavior. In L. A. Jeffress (Ed.), *Cerebral Mechanisms in Behavior: The Hixon Symposium* (pp. 112–146). Wiley.

Lee, Y. S., Ahn, S., Holt, R. F., & Schellenberg, E. G. (2020). Rhythm and Syntax Processing in School-Age Children. *Developmental Psychology, 56*(9). https://doi.org/10.1037/dev0000969

Lense, M. D., Gordon, R. L., Key, A. P. F., & Dykens, E. M. (2014). Neural correlates of cross-modal affective priming by music in williams syndrome. *Social Cognitive and Affective Neuroscience, 9*(4), 529–537. https://doi.org/10.1093/scan/nst017

Lerdahl, F., & Jackendoff, R. (1983). *A generative theory of tonal music*. MIT Press.

Leventhal, D. K., Gage, G. J., Schmidt, R., Pettibone, J. R., Case, A. C., & Berke, J. D. (2012). Basal ganglia beta oscillations accompany cue utilization. *Neuron, 73*(3), 523–536. https://doi.org/10.1016/j.neuron.2011.11.032

Luo, H., & Poeppel, D. (2007). Phase Patterns of Neuronal Responses Reliably Discriminate Speech in Human Auditory Cortex. *Neuron, 54*(6), 1001–1010. https://doi.org/10.1016/j.neuron.2007.06.004

Luo, H., Wang, Y., Poeppel, D., Simon, J. Z., & Ding, N. (2006). Concurrent Encoding of Frequency and Amplitude Modulation in Human Auditory Cortex: MEG Evidence. *Journal of Neurophysiology, 96*(5), 2712–2723. https://doi.org/10.1152/jn.01256.2005

Macmillan, N. A., & Creelman, C. D. (2005). *Detection Theory: A user's guide (Second)*. Mahwah, NJ, US: Lawrence Erlbaum Associates.

Maris, E., & Oostenveld, R. (2007). Nonparametric statistical testing of EEG- and MEG-data. *Journal of Neuroscience Methods, 164*(1), 177–190. https://doi.org/10.1016/j.jneumeth.2007.03.024

Mehler, J., Jusczyk, P., Lambertz, G., Halsted, N., Bertoncini, J., & Amiel-Tison, C. (1988). A precursor of language acquisition in young infants. *Cognition, 29*(2), 143–178. https://doi.org/10.1016/0010-0277(88)90035-2

Mehr, S. A., Singh, M., Knox, D., Ketter, D. M., Pickens-Jones, D., Atwood, S., et al. (2019). Universality and diversity in human song. *Science, 366*, 6468. https://doi.org/10.1126/science.aax0868

Morgan, J. L., Meier, R. P., & Newport, E. L. (1987). Structural packaging in the input to lanuage learning: Contributions of prosodic and morphological marking of phrases to the acquisition of language. *Cognitive Psychology, 19*(4), 498–550. https://doi.org/10.1016/0010-0285(87)90017-X

Nave-Blodgett, J. E., Snyder, J. S., & Hannon, E. E. (2021). Hierarchical beat perception develops throughout childhood and adolescence and is enhanced in those with musical training. *Journal of Experimental Psychology: General, 150*(2), 314–339. https://doi.org/10.1037/xge0000903

Nayak, S., Coleman, P. L., Ladányi, E., Nitin, R., Gustavson, D. E., Fisher, S. E., et al. (2022). The musical abilities, pleiotropy, language, and environment (MAPLE) framework for understanding musicality-language links across the lifespan. *Neurobiology of Language, 3*(4), 615–664. https://doi.org/10.1162/nol_a_00079

Niarchou, M., Gustavson, D. E., Sathirapongsasuti, J. F., Anglada-Tort, M., Eising, E., Bell, E., McArthur, E., Straub, P., The 23andMe Research Team, McAuley, J. D., Capra, J. A., Ullén, F., Creanza, N., Mosing, M. A., Hinds, D., Davis, L. K., Jacoby, N., & Gordon, R. L. (2022). Genome-wide association study of musical beat synchronization demonstrates high polygenicity. *Nature Human Behaviour, 6*(9), 1292-1309. Doi: 10.1038/s41562-022-01359-x.

Nitin, R., Gustavson, D. E., Aaron, A. S., Boorom, O. A., Bush, C. T., Wiens, N., et al. (2023). Exploring individual differences in musical rhythm and grammar skills in school-aged children with typically developing language. *Scientific Reports, 13*, 2201. https://doi.org/10.1038/s41598-022-21902-0

Obleser, J., & Kayser, C. (2019). Neural entrainment and attentional selection in the listening brain. *Trends in Cognitive Sciences, 23*(11), 913–926. https://doi.org/10.1016/j.tics.2019.08.004

Oostenveld, R., Fries, P., Maris, E., & Schoffelen, J. M. (2011). FieldTrip: Open source software for advanced analysis of MEG, EEG, and invasive electrophysiological data. *Computational Intelligence and Neuroscience, 2011*, Article 156869. https://doi.org/10.1155/2011/156869

Ozernov-Palchik, O., Wolf, M., & Patel, A. D. (2018). Relationships between early literacy and nonlinguistic rhythmic processes in kindergarteners. *Journal of Experimental Child Psychology, 167*, 354–368. https://doi.org/10.1016/j.jecp.2017.11.009

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*(May), 1–14. https://doi.org/10.3389/fnsys.2014.00057

Pearce, E., Launay, J., & Dunbar, R. I. M. (2015). The ice-breaker effect: Singing mediates fast social bonding. *Royal Society Open Science, 2*, Article 150221. https://doi.org/10.1098/rsos.150221

Perrin, F., Pernier, J., Bertrand, O., & Echallier, J. F. (1989). Spherical splines for scalp potential and current density mapping. *Electroencephalography and Clinical Neurophysiology, 72*(2), 184–187. https://doi.org/10.1016/0013-4694(89)90180-6

Persici, V., Stucchi, N., & Arosio, F. (2019). Predicting the future in rhythm and language: The anticipation abilities of a group of Italian-speaking children. In P. Guijarro-Fuentes & C. Suárez-Gómez (Eds.), Proceedings of GALA 2017: Language Acquisition and Development (pp. 451–468). Newcastle upon Tyne: Cambridge Scholar Publishers.

Pion-Tonachini, L., Kreutz-Delgado, K., & Makeig, S. (2019). ICLabel: An automated electroencephalographic independent component classifier, dataset, and website. *NeuroImage, 198*, 181–197. https://doi.org/10.1016/j.neuroimage.2019.05.026

Politimou, N., Dalla Bella, S., Farrugia, N., & Franco, F. (2019). Born to speak and sing: Musical predictors of language development in pre-schoolers. *Frontiers in Psychology, 10*(APR). https://doi.org/10.3389/fpsyg.2019.00948

Ponton, C. W., Eggermont, J. J., Kwong, B., & Don, M. (2000). Maturation of human central auditory system activity: Evidence from multi-channel evoked potentials. *Clinical Neurophysiology, 111*(2), 220–236. https://doi.org/10.1016/S1388-2457(99)00236-9

Przybylski, L., Bedoin, N., Krifi-Papoz, S., Herbillon, V., Roch, D., Léculier, L., et al. (2013). Rhythmic auditory stimulation influences syntactic processing in children with developmental language disorders. *Neuropsychology, 27*(1), 121–131. https://doi.org/10.1037/a0031277

R Core Team (2022). R: A Language and Environment for Statistical Computing. R Foundation for Statistical Computing, Vienna, Austria. URL: https://www.R-project.org/.

Radford, A. (1988). *Transformational Grammar: A First Course*. Cambridge, UK: Cambridge University Press.

Ravignani, A., Delgado, T., & Kirby, S. (2017). Musical evolution in the lab exhibits rhythmic universals. *Nature Human Behaviour, 1*, 0007. https://doi.org/10.1038/s41562-016-0007

Repp, B. H. (2005). Sensorimotor synchronization: A review of the tapping literature. *Psychonomic Bulletin and Review, 12*(6), 969–992. https://doi.org/10.3758/BF03206433

Repp, B. H., & Su, Y. H. (2013). Sensorimotor synchronization: A review of recent research (2006-2012). *Psychonomic Bulletin and Review, 20*(3), 403-452. Doi: 10.3758/s13423- 012-0371-2.

Roach, B. J., & Mathalon, D. H. (2008). Event-related EEG time-frequency analysis: An overview of measures and an analysis of early gamma band phase locking in schizophrenia. *Schizophrenia Bulletin, 34*(5), 907–926. https://doi.org/10.1093/schbul/sbn093

Salmelin, R., Hámáaláinen, M., Kajola, M., & Hari, R. (1995). Functional segregation of movement-related rhythmic activity in the human brain. *NeuroImage, 2*(4), 237–243.

Savage, P. E., Brown, S., Sakai, E., & Currie, T. E. (2015). Statistical universals reveal the structures and functions of human music. *Proceedings of the National Academy of Sciences, 112*(29), 8987–8992. https://doi.org/10.1073/pnas.1414495112

Savage, P. E., Loui, P., Tarr, B., Schachner, A., Glowacki, L., Mithen, S., et al. (2020). Music as a coevolved system for social bonding. *Behavioral and Brain Sciences, 44*, e59.

Schmidt-Kassow, M., & Kotz, S. A. (2008). Entrainment of syntactic processing? ERP-responses to predictable time intervals during syntactic reanalysis. *Brain Research, 1226*, 144–155. https://doi.org/10.1016/j.brainres.2008.06.017

Snyder, J. S., & Large, E. W. (2005). Gamma-band activity reflects the metric structure of rhythmic tone sequences. *Cognitive Brain Research, 24*(1), 117–126. https://doi.org/10.1016/j.cogbrainres.2004.12.014

Stupacher, J., Wood, G., & Witte, M. (2017). Neural entrainment to polyrhythms: A comparison of musicians and non-musicians. *Frontiers in Neuroscience, 11*(APR). https://doi.org/10.3389/fnins.2017.00208

Swaminathan, S., & Schellenberg, E. G. (2019). Musical ability, music training, and language ability in childhood. *Journal of Experimental Psychology: Learning Memory and Cognition, 46*(12), 2340–2348. https://doi.org/10.1037/xlm0000798

Tallon-Baudry, C., Bertrand, O., Delpuech, C., & Pernier, J. (1996). Stimulus specificity of phase-locked and non-phase-locked 40 Hz visual responses in human. *Journal of Neuroscience, 16*, 4240–4249. https://doi.org/10.1523/JNEUROSCI.16-13-04240.1996

Tarr, B., Launay, J., & Dunbar, R. I. M. (2016). Silent disco: Dancing in synchrony leads to elevated pain thresholds and social closeness. *Evolution and Human Behavior, 37*(5), 343–349. https://doi.org/10.1016/j.evolhumbehav.2016.02.004

The Mathworks Inc. (2017). MATLAB - MathWorks. *Www.Mathworks.Com/Products/Matlab*.

Vasilyeva, M., Waterfall, H., & Huttenlocher, J. (2008). Emergence of syntax: Commonalities and differences across children. *Developmental Science, 11*(1), 84–97. https://doi.org/10.1111/j.1467-7687.2007.00656.x

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

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 of the United States of America, 106*(7), 2468–2471. Doi: 10.1073/pnas.0809035106.

Woodruff Carr, K., White-Schwoch, T., Tierney, A. T., Strait, D. L., & Kraus, N. (2014). Beat synchronization predicts neural speech encoding and reading readiness in preschoolers. *Proceedings of the National Academy of Sciences of the United States of America, 111*(40), 14559–14564. Doi: 10.1073/pnas.1406219111.

Zanto, T. P., Large, E. W., Fuchs, A., & Kelso, J. A. S. (2005). Gamma-band responses to perturbed auditory sequences: Evidence for synchronization of perceptual processes. *Music Perception, 22*(3), 531–547. https://doi.org/10.1525/mp.2005.22.3.531