



OPEN

## Rhythmic motor behavior explains individual differences in grammar skills in adults

Hyun-Woong Kim<sup>1,2,3</sup>, Jessica Kovar<sup>1,2</sup>, Jesper Singh Bajwa<sup>1</sup>, Yasir Mian<sup>1</sup>, Ayesha Ahmad<sup>1,4</sup>, Marisol Mancilla Moreno<sup>1,4</sup>, Theodore J. Price<sup>1,4</sup> & Yune Sang Lee<sup>1,2,5</sup>✉

A growing body of literature has reported the relationship between music and language, particularly between individual differences in perceptual rhythm skill and grammar competency in children. Here, we investigated whether motoric aspects of rhythm processing—as measured by rhythmic finger tapping tasks—also explain the rhythm-grammar connection in 150 healthy young adults. We found that all expressive rhythm skills (spontaneous, synchronized, and continued tapping) along with rhythm discrimination skill significantly predicted receptive grammar skills on either auditory sentence comprehension or grammaticality well-formedness judgment (e.g., singular/plural, past/present), even after controlling for verbal working memory and music experience. Among these, synchronized tapping and rhythm discrimination explained unique variance of sentence comprehension and grammaticality judgment, respectively, indicating differential associations between different rhythm and grammar skills. Together, we demonstrate that even simple and repetitive motor behavior can account for seemingly high-order grammar skills in the adult population, suggesting that the sensorimotor system continue to support syntactic operations.

Rhythm is pervasive in a myriad of auditory events including music, language, and everyday listening sounds. Importantly, in the language domain, rhythmicity in continuous speech signals provides potent cues for tracking ongoing syntactic structures<sup>1,2</sup>. While previous theoretical frameworks situate language as a highly specialized faculty distinct from music<sup>3,4</sup>, a growing body of evidence demonstrates the relationship between musical rhythm and linguistic syntax<sup>5–8</sup>. For example, the ability to discriminate between pairs of short musical rhythms was positively associated with both receptive<sup>6,8</sup> and expressive<sup>5,7</sup> grammar proficiencies in typically developing children. Such rhythm-grammar connection may suggest co-optation of neurobiological resources shared by music and language<sup>9</sup>.

A core neuroanatomical substrate supporting rhythm processing is the sensorimotor system consisting of the basal ganglia, supplementary motor area, premotor cortex, and cerebellum. Indeed, synchronizing movements to musical rhythms is ubiquitous in human behaviors; listening to music often makes us spontaneously move to its rhythm. The sensorimotor brain regions are activated not only by rhythm production<sup>10–12</sup>, but also by passively listening to musical rhythms<sup>13–15</sup>. In particular, the basal ganglia and supplementary motor area have been shown to be sensitive to temporal regularity of musical rhythms that induces a sense of beat<sup>16–19</sup>. As such, it has been proposed that the motor circuitries connecting the basal ganglia to supplementary motor area support rhythm and beat perception by implicitly generating periodic actions towards upcoming beat timings<sup>20</sup>.

Given the crucial involvement of the sensorimotor system in rhythm perception, the motor component of rhythm processing may also explain individual differences in following rule-based temporal dynamics in the language domain, i.e., syntax. However, there are only a few studies that examined both beat synchronization and grammar skills in children<sup>21,22</sup>. In addition, although there are large individual differences in both motoric and perceptual rhythm skills in adults<sup>23,24</sup> as well as in children<sup>25</sup>, whether such differences can be translated to grammar skills in the adult population remains elusive. In the present study, we addressed this issue by measuring multiple rhythm and grammar skills in a large sample of healthy young adults.

Sensorimotor rhythmic skill is often assessed via synchronized finger tapping to an isochronous (i.e., metronome) tone sequence; the consistency (or variability) of tapping is used to measure the sensorimotor

<sup>1</sup>School of Behavioral and Brain Sciences, University of Texas at Dallas, Richardson, USA. <sup>2</sup>Callier Center for Communication Disorders, University of Texas at Dallas, Richardson, USA. <sup>3</sup>Department of Psychology, The University of Texas at Dallas, Richardson, USA. <sup>4</sup>Department of Neuroscience and Center for Advanced Pain Studies, University of Texas at Dallas, Richardson, USA. <sup>5</sup>Department of Speech, Language, and Hearing, The University of Texas at Dallas, Richardson, USA. ✉email: Yune.Lee@UTDallas.Edu

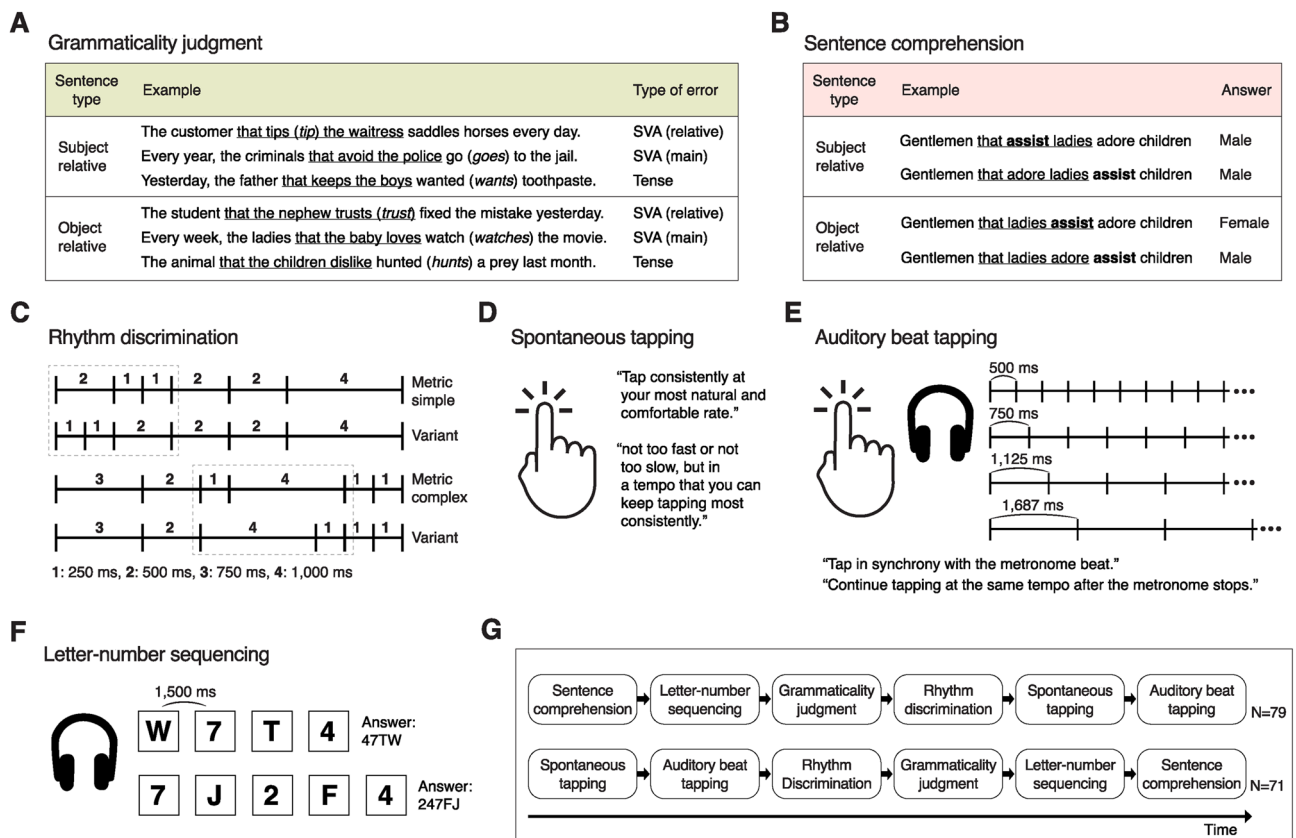
synchronization performance<sup>26</sup>. In addition, endogenous rhythmic behavior has been evaluated by means of spontaneous motor tapping that involves finger tapping at one's most natural and comfortable interval, often associated with 'personally preferred' or 'optimal' tempo<sup>27,28</sup>. Here we measured these motoric rhythm skills as well as a perceptual rhythm skill using a same/different rhythm discrimination task. For grammar measures, we employed two language tasks on spoken sentences wherein participants were asked to judge syntactic well-formedness of spoken sentences with either a subject- or object-relative center-embedded clause or to identify the grammatical agent linked to an action verb<sup>29,30</sup>. In addition, we assessed participants' verbal working memory, which was used as covariates along with demographic information including age, gender, and music background. We predicted that performance on the rhythm tasks would be associated with performance on the grammar tasks, even after controlling for verbal working memory.

## Results

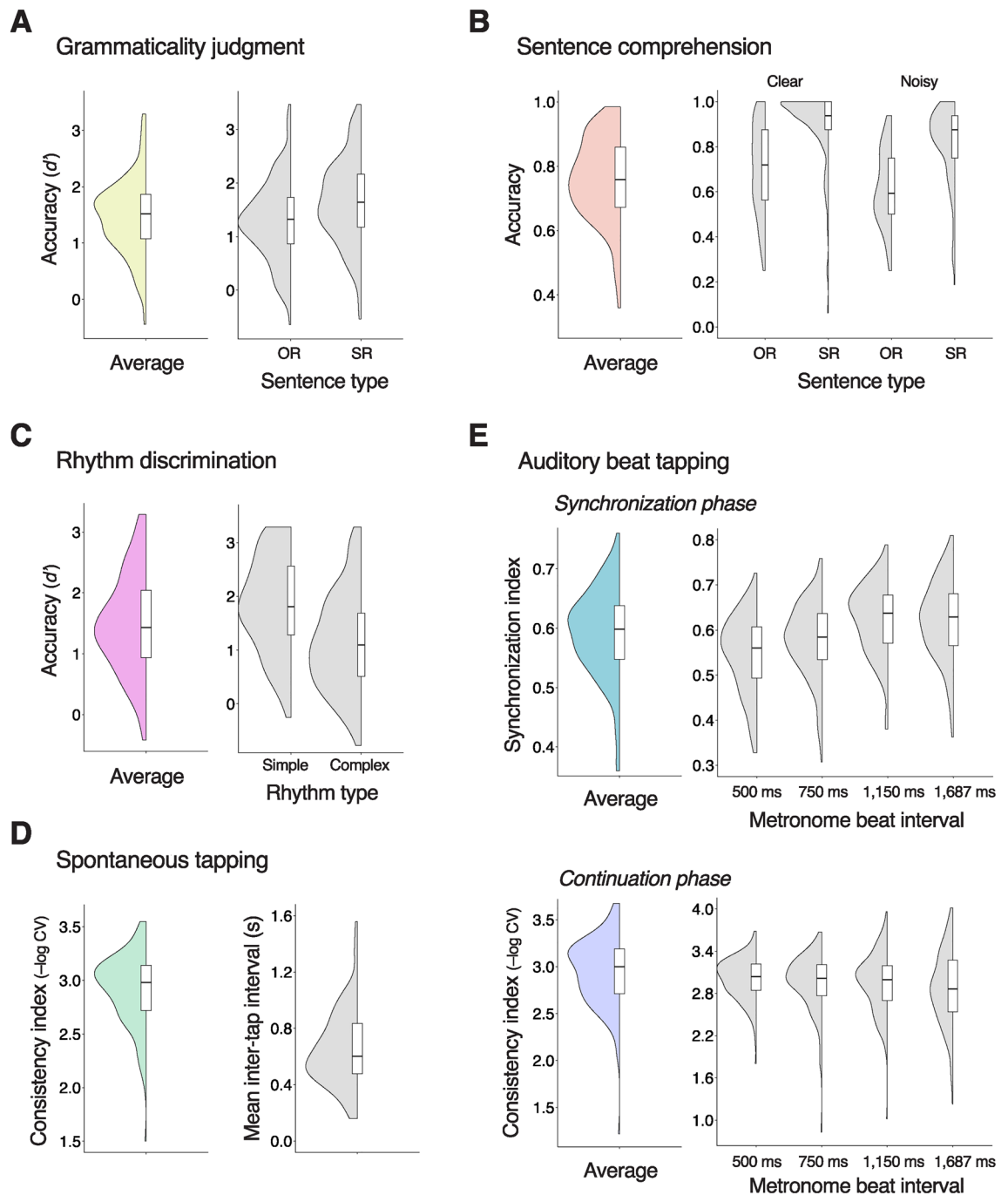
### Description of the behavioral tasks and outcomes

A total of 150 participants (mean age = 20.4 years, SD = 2.46) underwent two language tasks on spoken sentences (grammaticality judgment and sentence comprehension), three rhythm tasks (rhythm discrimination, spontaneous tapping, and auditory beat tapping), and a working memory task (letter-number sequencing) (Fig. 1). The behavioral results are summarized in Fig. 2 and Table 1.

The grammaticality judgment task assessed participants' ability to discern syntactic well-formedness on a series of spoken sentences that differed in syntactic complexity, containing either a subject-relative (SR) or object-relative (OR) center-embedded clause (Fig. 1A). Half of these sentences were grammatically correct while the



**Figure 1.** Overview of the experimental procedures. **(A)** Example sentences in grammaticality judgment task. Participants indicate if each spoken sentence is grammatically correct or not. Half of the sentences contain a subject-verb agreement (SVA) error or a past tense error. The relative clause is underlined, and the syntactic error is shown as italic in the parenthesis. **(B)** Example sentences in sentence comprehension task. Participants indicate the gender of individuals linked to an action verb, but not to four pre-designated preference verbs (love, adore, hate, and dislike), on each spoken sentence presented with or without a multi-talker babble noise. The relative clause is underlined, and the target action verb is in bold. **(C)** Schematic representation of rhythm sequences in rhythm discrimination task. Participants listen to each pair of rhythms and indicate if they were the same or different. **(D)** Spontaneous tapping task. Participants are instructed to tap consistently at their own tempo without external metronomes. **(E)** Auditory beat tapping task. Participants tap along with metronome beats presented in one of four tempos (inter-beat intervals of 500, 750, 1125, or 1687 ms) (synchronization phase) and continue tapping after the metronome stops (continuation phase). **(F)** Schematic representation of auditory sequence in letter-number sequencing task. Participants listen to a sequence of alternating letters and numbers and repeat them back in a sorted order. **(G)** Each participant completes the six behavioral tasks in one of the two orders shown. See Methods for more details.



**Figure 2.** Rhythm and grammar performances. Half violin plots display the distributions of task performance. Box plots display the median (horizontal line within the box), the first quantile (lower boundary), and the third quantile (upper boundary). (A) Grammaticality judgment accuracy. (B) Sentence comprehension accuracy. (C) Rhythm discrimination accuracy. (D) Spontaneous tapping consistency (left) and mean tapping interval (right). (E) Auditory beat tapping consistency during synchronization phase (top) and continuation phase (bottom).

other half had a morpho-syntactic error. As expected, participants performed worse on detecting grammatical errors in OR compared to SR sentences (Fig. 2A;  $F(1,149) = 38.1$ ,  $P = 6 \times 10^{-9}$ ), due to its non-canonical structure<sup>31</sup>.

The sentence comprehension task also consisted of a series of SR and OR sentences (all syntactically well-formed). Half of these sentences were mixed with multi-talker babble noise. In this task, participants identified the gender of a noun (e.g., boys) linked to an action verb (e.g., assist), while ignoring a noun related to any of the following preference verbs: love, adore, dislike, and hate (Fig. 1B). As was the case with the grammaticality judgment task, OR sentences yielded lower comprehension accuracy than SR sentences ( $\chi^2(1) = 527.8$ ,  $P < 2 \times 10^{-16}$ ). The background noise yielded lower accuracy ( $\chi^2(1) = 125.7$ ,  $P < 2 \times 10^{-16}$ ), especially for OR sentences, resulting in significant interaction with the sentence type ( $\chi^2(1) = 10.7$ ,  $P = 0.001$ ) (Fig. 2B).

Variable	<i>M</i>	<i>SD</i>	Median	Min	Max
Gender (F:M)	80:70	–	–	–	–
Age (years)	20.5	2.49	20	18	37
Years of musical training	4.24	4.69	3	0	17
Grammaticality judgment accuracy ( <i>d'</i> )	1.47	0.66	1.52	–0.44	3.67
Sentence comprehension accuracy (acc)	0.76	0.13	0.76	0.36	0.98
Rhythm discrimination accuracy ( <i>d'</i> )	1.49	0.80	1.43	–0.42	3.29
Spontaneous tapping consistency (–log CV)	2.91	0.33	2.98	1.5	3.55
Beat tapping consistency: synchronization ( <i>SI</i> )	0.59	0.07	0.60	0.36	0.76
Beat tapping consistency: continuation (–log CV)	2.94	0.36	3	1.22	3.67
Letter-number sequencing accuracy (N correct)	12.8	3.44	12	4	24
Spontaneous tapping mean interval (s)	0.66	0.264	0.602	0.160	1.558

**Table 1.** Mean raw scores (*M*) and standard deviations (*SD*) of demographic and behavioral variables.

In the rhythm discrimination task, participants determined whether pairs of short rhythmic sequences were the same or different (Fig. 1C). Half of the rhythms were metrically simple in that the inter-tone intervals were regularly arranged into groups of 1-s beat intervals, rendering a sense of beat<sup>16</sup>. The other half of rhythms were metrically complex, making it difficult to detect the underlying beat. Consistent with previous findings<sup>16,32</sup>, metrically complex rhythms were more difficult to discriminate than simple rhythms (Fig. 2C;  $F(1,149) = 105.6$ ,  $P < 2 \times 10^{-16}$ ).

In the spontaneous tapping task, participants were instructed to tap their right index finger consistently at their most natural and comfortable tempo (Fig. 1D). The average tempo (0.66 s) and its range (0.18–1.6 s) were similar to those in the previous report<sup>28</sup>. We gauged participants' internal rhythm performance using a consistency index derived from the variance of inter-tap intervals (ITI; see Methods for more details).

In the auditory beat tapping task, participants tapped their right index finger to metronome beats with inter-beat intervals (IBIs) of 500, 750, 1125, and 1687 ms (synchronization phase) (Fig. 1E). These IBIs roughly correspond to 120, 80, 53, and 36 beats-per-minutes (BPM). Participants were also required to continue tapping consistently at the same tempo after the metronome stopped halfway through (continuation phase). A synchronization index was calculated from the uncertainty of the distribution of tap timings with respect to beat timings for the synchronization data, and a consistency index was computed from the ITI variance for the continuation data (see Methods for more details). We found that longer IBIs yielded more synchronized tapping responses during the synchronization phase ( $\chi^2(3) = 177.0$ ,  $P < 2 \times 10^{-16}$ ), while shorter IBIs yielded more consistent tapping during the continuation phase ( $\chi^2(3) = 16.4$ ,  $P = 9 \times 10^{-4}$ ) (Fig. 2E).

Lastly, we measured participants' working memory using the letter-number sequencing task (Fig. 1F). In this task, participants were required to verbally re-organize a series of alternating numbers and letters in ascending orders (e.g., for K3F7R5, the correct answer is 357FKR). Working memory has been shown to be positively associated with a wide range of music and language abilities including sentence comprehension<sup>33</sup> and grammaticality judgment<sup>34</sup>, as well as timing<sup>35</sup>, perceptual rhythm<sup>25</sup>, and motor rhythm<sup>36</sup>. Thus, we included the working memory measure as a covariate along with demographic variables of age, gender, and music background in the regression analyses.

### Multiple regression analysis

The regression results are listed in Table 2. The *P* values were adjusted with the false discovery rate (FDR) correction for multiple testing. We found that rhythm discrimination was significantly associated with grammaticality judgment ( $b = 0.24$ ,  $t = 2.95$ ,  $P = 0.009$ , FDR corrected), but not with sentence comprehension ( $b = 0.16$ ,  $t = 1.93$ ,  $P = 0.064$ , FDR corrected). By contrast, spontaneous tapping consistency was associated with both grammaticality judgment ( $b = 0.23$ ,  $t = 2.73$ ,  $P = 0.012$ , FDR corrected) and sentence comprehension ( $b = 0.26$ ,  $t = 3.21$ ,  $P = 0.006$ , FDR corrected). The strongest rhythm-grammar association among all pairwise regressions emerged between beat synchronization tapping and sentence comprehension ( $b = 0.40$ ,  $t = 5.05$ ,  $P = 1 \times 10^{-6}$ , FDR corrected), while this rhythm skill was not predictive of grammaticality judgment ( $b = 0.14$ ,  $t = 1.63$ ,  $P = 0.106$ , FDR corrected). Lastly, continuation tapping was related to both grammaticality judgment ( $b = 0.19$ ,  $t = 2.11$ ,  $P = 0.049$ , FDR corrected) and sentence comprehension ( $b = 0.26$ ,  $t = 2.99$ ,  $P = 0.009$ , FDR corrected).

To further learn whether different rhythm skills account for unique variance in grammar skills, we performed multiple regression analyses with all four rhythm measures—that were highly inter-related with each other—included as predictors. In this most stringent model, beat synchronization still robustly accounted for sentence comprehension ( $b = 0.40$ ,  $t = 3.62$ ,  $P = 4 \times 10^{-4}$ ), whereas rhythm discrimination ( $b = 0.20$ ,  $t = 2.22$ ,  $p = 0.028$ ) was predictive of grammaticality judgment, even after controlling for the other rhythm measures (Table 3). This suggests that there are dissociable aspects of temporal processing in the relationship between rhythm and grammar skills.

Predictors	Grammaticality judgment			Sentence comprehension		
	<i>b</i>	<i>t</i>	<i>p</i>	<i>b</i>	<i>t</i>	<i>p</i>
Rhythm discrimination	<b>0.24</b>	<b>2.95</b>	<b>0.004</b>	0.16	1.93	0.056
<i>Verbal working memory</i>	<i>0.31</i>	<i>3.77</i>	<i>&lt;0.001</i>	<i>0.30</i>	<i>3.57</i>	<i>&lt;0.001</i>
<i>Years of musical training</i>	<i>0.11</i>	<i>1.52</i>	<i>0.131</i>	<i>0.18</i>	<i>2.44</i>	<i>0.016</i>
<i>Age</i>	<i>0.04</i>	<i>0.51</i>	<i>0.610</i>	<i>-0.01</i>	<i>-0.18</i>	<i>0.861</i>
<i>Gender (Male)</i>	<i>-0.01</i>	<i>-0.17</i>	<i>0.865</i>	<i>0.10</i>	<i>1.27</i>	<i>0.205</i>
Spontaneous tapping	<b>0.23</b>	<b>2.73</b>	<b>0.007</b>	<b>0.26</b>	<b>3.22</b>	<b>0.002</b>
<i>Verbal working memory</i>	<i>0.35</i>	<i>4.59</i>	<i>&lt;0.001</i>	<i>0.30</i>	<i>3.96</i>	<i>&lt;0.001</i>
<i>Years of musical training</i>	<i>0.07</i>	<i>0.82</i>	<i>0.411</i>	<i>0.11</i>	<i>1.37</i>	<i>0.172</i>
<i>Age</i>	<i>0.01</i>	<i>0.11</i>	<i>0.915</i>	<i>-0.05</i>	<i>-0.62</i>	<i>0.539</i>
<i>Gender (Male)</i>	<i>-0.04</i>	<i>-0.59</i>	<i>0.555</i>	<i>0.07</i>	<i>0.87</i>	<i>0.385</i>
Beat tapping: Synchronization	0.14	1.63	0.110	<b>0.40</b>	<b>5.05</b>	<b>&lt;0.001</b>
<i>Verbal working memory</i>	<i>0.38</i>	<i>4.80</i>	<i>&lt;0.001</i>	<i>0.28</i>	<i>3.84</i>	<i>&lt;0.001</i>
<i>Years of musical training</i>	<i>0.10</i>	<i>1.16</i>	<i>0.250</i>	<i>0.04</i>	<i>0.56</i>	<i>0.576</i>
<i>Age</i>	<i>0.02</i>	<i>0.30</i>	<i>0.764</i>	<i>-0.05</i>	<i>-0.65</i>	<i>0.516</i>
<i>Gender (Male)</i>	<i>-0.05</i>	<i>-0.62</i>	<i>0.539</i>	<i>0.02</i>	<i>0.27</i>	<i>0.790</i>
Beat tapping: Continuation	<b>0.19</b>	<b>2.11</b>	<b>0.037</b>	<b>0.26</b>	<b>2.99</b>	<b>0.003</b>
<i>Verbal working memory</i>	<i>0.35</i>	<i>4.31</i>	<i>&lt;0.001</i>	<i>0.28</i>	<i>3.54</i>	<i>&lt;0.001</i>
<i>Years of musical training</i>	<i>0.08</i>	<i>1.04</i>	<i>0.301</i>	<i>0.11</i>	<i>1.43</i>	<i>0.156</i>
<i>Age</i>	<i>0.02</i>	<i>0.21</i>	<i>0.836</i>	<i>-0.04</i>	<i>-0.55</i>	<i>0.585</i>
<i>Gender (Male)</i>	<i>-0.06</i>	<i>-0.79</i>	<i>0.434</i>	<i>0.04</i>	<i>0.48</i>	<i>0.630</i>

**Table 2.** Standardized regression coefficients, *t* values, and uncorrected *P* values in predicting each of two grammar measures as a function of demographic variables and each of four rhythm measures. Significant values are in bold, and covariates are italicized.

Predictors	Grammaticality judgment			Sentence comprehension		
	<i>b</i>	<i>t</i>	<i>p</i>	<i>b</i>	<i>t</i>	<i>p</i>
Rhythm discrimination	<b>0.20</b>	<b>2.22</b>	<b>0.028</b>	0.02	0.21	0.834
Spontaneous tapping	0.19	1.68	0.095	0.11	0.98	0.327
Beat tapping: Synchronization	-0.03	-0.03	0.764	<b>0.40</b>	<b>3.62</b>	<b>&lt;0.001</b>
Beat tapping: Continuation	-0.01	-0.10	0.918	-0.09	-0.70	0.483
<i>Working memory</i>	<i>0.29</i>	<i>3.55</i>	<i>&lt;0.001</i>	<i>0.28</i>	<i>3.51</i>	<i>&lt;0.001</i>
<i>Years of musical training</i>	<i>0.07</i>	<i>0.79</i>	<i>0.429</i>	<i>0.04</i>	<i>0.46</i>	<i>0.648</i>
<i>Age</i>	<i>0.02</i>	<i>0.26</i>	<i>0.792</i>	<i>-0.05</i>	<i>-0.67</i>	<i>0.502</i>
<i>Gender (Male)</i>	<i>-0.02</i>	<i>-0.30</i>	<i>0.764</i>	<i>0.03</i>	<i>0.40</i>	<i>0.688</i>

**Table 3.** Standardized regression coefficients, *t* values, and uncorrected *P* values in predicting each of two grammar measures as a function of demographic variables and all four rhythm measures. Significant values are in bold, and covariates are italicized.

## Discussion

Although a growing body of research has demonstrated connections between musical rhythm and linguistic grammar skills<sup>5–8</sup>, the evidence has been limited in children and mostly to a perceptual rhythm skill, i.e., musical rhythm discrimination. In the present study with 150 healthy young adults, we demonstrated that a set of rhythm skills in both receptive and expressive domains were associated with receptive grammar tasks. Notably, even simple and repetitive motor behavior involving spontaneous or synchronized finger tapping was predictive of comprehension on spoken sentences that varied in syntactic structure. This finding extends the existing evidence beyond the perceptual rhythm toward expressive/motoric rhythm skills, as well as from children to young adults. Together, these results suggest that common neurobiological mechanisms may be at play in both rhythm and syntactic processing<sup>9</sup>, contributing to the association between individual differences in rhythm and grammar that persists into adulthood.

The present findings may shed light on the role of the motor system in auditory syntactic processing<sup>37</sup>, by showing that even a relatively simple motor task such as spontaneous or continued finger tapping explained both grammaticality judgment and syntactic interpretation on spoken sentences. Spontaneous rhythmic behavior has been theorized to reflect activity of self-sustaining internal oscillators<sup>27,28</sup>, which is likely regulated by

cortico-striatal motor circuits<sup>38,39</sup>. Our data speak to the functional role of internal rhythmicity in readily analyzing syntactic structures during auditory sentence processing. In addition, beat synchronization tapping had a robust explanatory power in predicting sentence comprehension performance, which was significant even when controlling for the effects of the other rhythm measures. Given that sensorimotor synchronization to external rhythm is thought to rely on predictive timing mechanisms<sup>26,40</sup>, the current finding supports the idea that temporal prediction (e.g., when the next event occurs) may play a functional role in syntactic prediction (e.g., what word comes next given the preceding words in a sentence)<sup>37,41</sup>.

By contrast, the synchronized tapping measurements had a less robust relationship with grammaticality judgement performance, suggesting that this task may require some different mechanisms that cannot be solely explained by (forward) temporal prediction, such as (backward) re-analysis of preceding syntactic contents<sup>42,43</sup>. The re-analysis of temporal structure may have been captured by the significant relationship between rhythm discrimination and grammaticality judgment. This is perhaps because both tasks require back-and-forth comparison of words or tones in relation to the preceding ones to achieve timely judgement of the linguistic or rhythmic structure. This finding is in line with the idea that musical rhythm processing may recruit neurobiological resources for rule-based temporal processing shared by syntax processing system<sup>9</sup>.

Together, the current results suggest that different rhythm skills may uniquely contribute to accounting for different grammar skills. This underscores the importance of assessing multiple rhythm skills to gain a more complete picture of the rhythm-grammar relationship, which may be manifested through multiple neurocognitive mechanisms such as precise temporal predictions and re-analysis of temporal structures. Moreover, from a broader perspective, our findings suggest that the relationship between music and language may not be a unitary construct, but rather a consequence of multiple mechanisms shared by different aspects of music and language. For instance, phonological processing may be correlated more dominantly with components of rhythm processing that tap into precise auditory encoding<sup>44,45</sup>. Thus, future investigation of a wide range of music and language tasks using a well-powered sample may provide a more comprehensive understanding of the multifaceted associations between music and language.

In sum, the current study provides evidence for the association between expressive rhythm skills and receptive grammar skills in an adult population. This is in line with clinical observations that children with developmental language disorder often exhibit deficits in producing rhythmic movements<sup>46</sup>, which might build upon common genetic underpinnings<sup>47,48</sup>. Moreover, we found different rhythm skills uniquely explained different grammar skills, suggesting that there are dissociable aspects of temporal processing in the rhythm-grammar relationship. Our findings suggest that what has been regarded as a core linguistic operation, i.e., syntax, is associated with domain-general temporal processing in the sensorimotor system.

## Methods

### Participants

150 participants were recruited from the University of Texas at Dallas (80 females, 18–37 years, mean age = 20.4 years,  $SD = 2.5$  years). An a priori power analysis was performed using G × Power version 3.1, with an alpha level of  $0.05 \times 8 = 0.0063$ , considering that there were eight pairwise multiple regression tests to be performed (see below). The result showed that at least 118 participants were required to find a medium effect size ( $f^2 = 0.15$ ,  $\alpha$  error = 0.0063, Power = 0.95) with five predictors in a multiple regression analysis.

All participants spoke English as a primary language, had normal vision and hearing, and had no self-reported history of cognitive, developmental, or neurological disorders. All participants provided written informed consent to participate in the study. Participants' musical experience was quantified as the number of years of actively practicing one or multiple musical instruments including voice (Table 1). The experimental protocols were approved by the University of Texas at Dallas Institutional Review Board (IRB-21-109) and conducted in accordance with the Declaration of Helsinki. All methods were carried out in agreement with the relevant guidelines and regulations.

### Experimental procedures

The experiment consisted of the following behavioral tasks: two for grammar (grammaticality judgment, sentence comprehension); three for rhythm (rhythm discrimination, spontaneous tapping, auditory beat tapping); one for verbal working memory (letter-number sequencing) (Figs. 1, 2). Each participant completed the tasks in either of two fixed orders, as shown in Fig. 1G. The experimental procedures were conducted using Matlab R2021 (Mathworks, MA) in a dimly lit sound-proof booth. All auditory stimuli were presented at a preset volume (70 dB SPL) through Sennheiser HD-280 headphones. We used Google Text-to-Speech to generate spoken sentence stimuli for the grammar tasks and verbal items for the working memory task. The speaker voice was set to an American-English speaking male for the language tasks and a female for the working memory task.

#### *Grammaticality judgment task*

The language materials consisted of 48 sentences, each of which contained either a subject-relative (SR) or object-relative (OR) center-embedded clause and a time adverb phrase (e.g., every week, last year) (Fig. 1A). Half of the sentences contained a morpho-syntactic error of one of three types: a subject-verb agreement error within the relative phrase, a subject-verb agreement error in the main phrase, or a tense error (Fig. 1A). The sentence type (SR and OR) and the type of error were counterbalanced across trials. In each trial, participants listened to a spoken sentence and judged whether the sentence was grammatically correct or not via button press (right arrow for 'grammatical' and left arrow for 'ungrammatical'). Participants completed 8 practice trials, using sentences reserved only for practice (these sentences did not appear in the main task), with no visual presentation of the sentences. They received visual feedback after each response ('Correct' or 'Incorrect') and were able

to immediately repeat the practice trial if they chose. There was no time constraint in this task. Participants underwent 48 trials presented in a randomized order without feedback during the main task. A 15-s break was provided every 12 trials. We computed a  $d'$ -prime ( $d'$ ) score for each sentence type, which was obtained by calculating the difference of the  $z$ -scores between hit and false alarm rates. The rates of 0 and 1 were adjusted to prevent an indefinite  $d'$  score<sup>49</sup>.

#### *Sentence comprehension task*

The language materials were comprised of 64 base sentences, each consisting of six words: a male noun (e.g., men, sons, kings, etc.), a female noun (e.g., women, daughters, queens, etc.), a gender-neutral noun (e.g., children, farmers, artists, etc.), a relative pronoun 'that', a transitive action verb (e.g., help, protect, tease, bully, etc.), and one of four transitive preference verbs: love, adore, hate, and dislike. The six words were arranged to contain either an SR or OR center-embedded structure by switching the position of the noun and the verb within the relative clause (Fig. 1B). An action verb and a preference verb were located in either the main or relative clause. Half of the sentences were mixed with a background noise of multi-talker babble consisting of two male and two female speakers at a signal-to-noise ratio of  $-1$  dB. The sentence type (i.e., SR or OR), the clarity of the sentence (i.e., clear or noisy), the action verb location (i.e., main or relative clause), and the gender of the agent (i.e., female or male) were fully counterbalanced across trials. In each trial, participants listened to a spoken sentence and indicated the gender of the individuals performing an action, while disregarding those who love/adore/hate/dislike others, by pressing either the 'male' (left arrow) or 'female' (right arrow) key within 3 s. After the task instruction, participants underwent 8 practice trials during which the corresponding sentences were concurrently presented on the screen. They were provided with visual feedback and (if needed) additional verbal instruction from the experimenter following each response. Then they received 16 more trials with visual feedback but without visual presentation of the sentences. During the main task, participants completed a total of 64 trials presented in a randomized order without feedback and without visual presentation of the sentences. The sentences used during the practice trials were not presented. There was a 15-s break after every 16 trials.

Percent accuracy for each combination of sentence type and clarity conditions is displayed in Fig. 2B. We used a mixed effects logistic regression model (glmer in the *lme4* package)<sup>50</sup> to analyze the binary correct/incorrect responses, with the factors of sentence type (SR and OR), clarity (clear and noisy), and an interaction between the two as fixed effects and with a random intercept of participant. We evaluated statistical significance using the Type III Wald chi-square tests in the *car* package<sup>51</sup>.

#### *Rhythm discrimination task*

Rhythm stimuli were comprised of 10 metrically simple and 10 metrically complex rhythmic tone sequences and their respective variants, chosen from a set of rhythms used in Grahn and Brett<sup>16</sup>. Each rhythm consisted of seven or eight woodblock sounds (20 ms in duration) with intervals of 250, 500, 750, or 1000 ms between the sounds, spanning 3 s. A variant of each rhythm was made by switching the order of two adjacent intervals in the rhythm (Fig. 1C). On each trial, participants listened to a pair of rhythms presented sequentially with a 2-s interval and judged whether the two rhythms were the same or different by pressing either the 'same' (left arrow) or 'different' (right arrow) key. There was no time constraint in this task. Each rhythm was presented twice, once paired with itself (i.e., 'same') and once paired with its variant (i.e., 'different'), resulting in a total of 40 trials presented in a randomized order. Participants performed two practice trials (one 'same' and one 'different') with feedback using one of the 10 simple rhythms and its variant. Participants were allowed to replay the practice if they chose and hear the same two rhythm pairs again. No feedback was given during the main task. We computed a  $d'$  score for each rhythm type. The  $d'$  data were analyzed using a repeated-measures ANOVA with a within-subject factor of rhythm type (metric simple and metric complex).

#### *Spontaneous tapping task*

In the spontaneous tapping task, participants were instructed to tap the space bar with their right index finger consistently at their most natural and comfortable rate. They were encouraged not to tap too fast or too slow but at a tempo that they could keep tapping most consistently (Fig. 1D). Each of the four trials was terminated after 43 taps on the space bar (i.e., 42 ITIs). There was a short practice with 15 taps. The task was performed prior to the auditory beat tapping task (Fig. 1G), to prevent influence of exposure to metronome beats on spontaneous motor tempo. The first two ITIs were discarded from each trial. We obtained a mean ITI from the remaining 40 ITIs, which was calculated after excluding ITIs exceeding the initial mean ITI  $\pm 3$  standard deviations. We obtained a consistency index from a coefficient of variation (CV) of ITIs, i.e., a standard deviation of the remaining ITIs divided by the mean ITI. The consistency index was computed as  $-\log(\text{CV})$  for each trial and averaged across the four trials (Fig. 2D).

#### *Auditory beat tapping task*

Each trial of the task started with synchronization and ended with continuation. During the former phase, participants tapped the space bar with their right index finger in synchrony with a metronome beat (i.e., isochronous woodblock sounds each with 20 ms in duration). They were instructed to start tapping after the first four beats, accompanied by a countdown timer on the screen. The metronome stopped immediately after 21 taps (i.e., 20 ITIs), after which the continuation phase began wherein participants continued tapping consistently while trying to maintain the same tempo without metronome beats. The continuation phase lasted until receiving 21 tapping responses. Each trial used one of four IBIs, equally spaced in log scale: 500, 750, 1125, and 1687 ms (Fig. 1E). Each of these four tempos was repeated four times for a total of 16 trials. The presentation order of the four tempos was pseudo-randomized within each four-trial set, such that there was no transition from shortest to longest

IBI or vice versa. Participants underwent a practice trial with an IBI of 850 ms and were allowed to repeat the practice until understanding the instruction. A 30-s break was given after each set.

We analyzed beat tapping data during the synchronization and continuation phases separately. For the synchronization tapping data, we computed a circular metric (*SI*) based on Shannon entropy (*SE*):

$$SE = - \sum_{i=1}^M p(i) \ln p(i), \quad SI = 1 - \frac{SE}{\ln N}$$

where *M* is the number of bins, covering the range from  $-180^\circ$  to  $+180^\circ$ ; *p*(*i*) is the probability of tap timings (i.e., relative phase angles with respect to nearest beat timings) assigned to *i*<sub>th</sub> bin; *N* is the total number of tapping responses. The bin size was set to  $15^\circ$  based on an optimal bin search procedure described in Kim et al.<sup>25</sup> The *SE* quantifies the degree of spreading of a data distribution, i.e., the relative phase distribution in our data. We derived a synchronization index (*SI*) from *SE* as shown in the formula above. *SI* ranges from 0 (all responses occurred in different bins) to 1 (all responses occurred in a single bin), exhibiting a less skewed distribution compared to other consistency measures (e.g.,<sup>23</sup>). For the continuation tapping data with no reference beat timings, we computed a consistency index in the same way as for the spontaneous tapping data. The synchronization and consistency indices were computed for each tapping trial and averaged across the four trials for each tempo condition. The synchronization and consistency indices were analyzed using the Friedman test, a nonparametric test for repeated-measures data, with a within-subject factor of IBI (500, 750, 1125, and 1687 ms).

#### Letter-number sequencing task

The verbal items consisted of a set of spoken letters ('C', 'D', 'F', 'H', 'J', 'K', 'L', 'P', 'Q', 'R', 'S', 'T', 'W') and numbers (from 'one' to 'nine'). On each trial, participants heard a list of alternating numbers and letters presented via headphones with a 1.5-s interval (Fig. 1F). After presentation of the sequence, they were required to say the numbers in order from the smallest to the largest and then the letters in alphabetical order. An experimenter manually recorded the correctness. There were two practice trials with a list length of three and four. The main task started from a set of four trials with three list items each, and the list length increased after every four trials until participants failed to recall all trials in each length, after which the experiment ended. The largest length was eight, resulting in a maximum of 24 trials. The letter-number sequencing score was computed as the number of trials correctly recalled.

#### Multiple linear regression analysis

For each of the behavioral tasks, we used a representative outcome by collapsing scores across task conditions (Fig. 2). For example, a single measure was computed for synchronization by averaging consistency measures (*SI*) across four IBI conditions. A multiple linear regression analysis was performed for each combination of the two grammar and four rhythm measures, in which one of the four rhythm measures, working memory (i.e., letter-number sequencing score), gender, age, and years of musical training were entered as independent variables into the model to predict each of the two grammar measures (Table 2). The subsequent *P* values of the regression coefficients were adjusted with the false discovery rate (FDR) procedure ( $\alpha = 0.05$ ) to correct for multiple testing.

#### Data availability

The datasets generated and analyzed during the current study are available from the corresponding author on reasonable request.

Received: 6 June 2023; Accepted: 31 January 2024

Published online: 14 February 2024

#### References

- Nakamura, C., Arai, M. & Mazuka, R. Immediate use of prosody and context in predicting a syntactic structure. *Cognition* **125**, 317–323 (2012).
- Steinhauer, K., Alter, K. & Friederici, A. D. Brain potentials indicate immediate use of prosodic cues in natural speech processing. *Nat. Neurosci.* **2**, 191–196 (1999).
- Chomsky, N. *Aspects of the Theory of Syntax* (MIT Press, 1965).
- Pinker, S. & Jackendoff, R. The faculty of language: What's special about it?. *Cognition* **95**, 201–236 (2005).
- Gordon, R. L. et al. Musical rhythm discrimination explains individual differences in grammar skills in children. *Dev. Sci.* **18**(4), 635 (2015).
- Lee, Y. S., Ahn, S., Holt, R. F. & Schellenberg, E. G. Rhythm and syntax processing in school-age children. *Dev. Psychol.* <https://doi.org/10.1037/dev0000969> (2020).
- Nitin, R. et al. Exploring individual differences in musical rhythm and grammar skills in school-aged children with typically developing language. *Sci. Rep.* **13**, 2201 (2023).
- Swaminathan, S. & Schellenberg, E. G. Musical ability, music training, and language ability in childhood. *J. Exp. Psychol. Learn. Mem. Cognit.* <https://doi.org/10.1037/xlm0000798> (2019).
- Heard, M. & Lee, Y. S. Shared neural resources of rhythm and syntax: An ALE meta-analysis. *Neuropsychologia* **137**, 107284 (2020).
- Jantzen, K. J., Oullier, O., Marshall, M., Steinberg, F. L. & Kelso, J. A. S. A parametric fMRI investigation of context effects in sensorimotor timing and coordination. *Neuropsychologia* **45**, 673–684 (2007).
- Joliot, M. et al. Anatomical congruence of metabolic and electromagnetic activation signals during a self-paced motor task: A combined PET–MEG study. *NeuroImage* **7**, 337–351 (1998).
- Lewis, P. A., Wing, A. M., Pope, P. A., Praamstra, P. & Miall, R. C. Brain activity correlates differentially with increasing temporal complexity of rhythms during initialisation, synchronisation, and continuation phases of paced finger tapping. *Neuropsychologia* **42**, 1301–1312 (2004).



13. Chen, J. L., Penhune, V. B. & Zatorre, R. J. Listening to musical rhythms recruits motor regions of the brain. *Cereb. Cortex* **18**, 2844–2854 (2008).
14. Thaut, M., Trimarchi, P. & Parsons, L. Human brain basis of musical rhythm perception: Common and distinct neural substrates for meter, tempo, and pattern. *Brain Sci.* **4**, 428–452 (2014).
15. Bengtsson, S. L. *et al.* Listening to rhythms activates motor and premotor cortices. *Cortex* **45**, 62–71 (2009).
16. Grahn, J. A. & Brett, M. Rhythm and beat perception in motor areas of the brain. *J. Cogn. Neurosci.* **19**, 893–906 (2007).
17. Grahn, J. A. & Rowe, J. B. Feeling the beat: Premotor and striatal interactions in musicians and nonmusicians during beat perception. *J. Neurosci.* **29**, 7540–7548 (2009).
18. Grahn, J. A. & Rowe, J. B. Finding and feeling the musical beat: Striatal dissociations between detection and prediction of regularity. *Cereb. Cortex* **23**, 913–921 (2013).
19. Toiviainen, P., Burunat, I., Brattico, E., Vuust, P. & Alluri, V. The chronnectome of musical beat. *NeuroImage* **216**, 116191 (2020).
20. Cannon, J. J. & Patel, A. D. How beat perception co-opts motor neurophysiology. *Trends Cognit. Sci.* **25**, 137–150 (2021).
21. Cohrdes, C., Grolig, L. & Schroeder, S. Relating language and music skills in young children: A first approach to systemize and compare distinct competencies on different levels. *Front. Psychol.* **7**, 1616 (2016).
22. Politimou, N., Dalla Bella, S., Farrugia, N. & Franco, F. Born to speak and sing: Musical predictors of language development in pre-schoolers. *Front. Psychol.* **10**, 948 (2019).
23. Fujii, S. & Schlaug, G. The Harvard Beat Assessment Test (H-BAT): A battery for assessing beat perception and production and their dissociation. *Front. Hum. Neurosci.* **7**, 771 (2013).
24. Dalla Bella, S. *et al.* BAASTA: Battery for the assessment of auditory sensorimotor and timing abilities. *Behav. Res. Methods* **49**, 1128–1145 (2017).
25. Kim, H. W., Lee, K. M. & Lee, Y. S. Sensorimotor and working memory systems jointly support development of perceptual rhythm processing. *Dev. Sci.* **26**(1), e13261 (2023).
26. Repp, B. H. Sensorimotor synchronization: A review of the tapping literature. *Psychon. Bull. Rev.* **12**, 969–992 (2005).
27. Klinger Amrani, A. & Zion Golombic, E. Spontaneous and stimulus-driven rhythmic behaviors in ADHD adults and controls. *Neuropsychologia* **146**, 107544 (2020).
28. McAuley, J. D., Jones, M. R., Holub, S., Johnston, H. M. & Miller, N. S. The time of our lives: Life span development of timing and event tracking. *J. Exp. Psychol. Gen.* **135**, 348–367 (2006).
29. Kim, H. W., McLaren, K. E. & Lee, Y. S. No influence of regular rhythmic priming on grammaticality judgment and sentence comprehension in English-speaking children. *J. Exp. Child Psychol.* **237**, 105760 (2024).
30. Kim, H. W., Happe, J. & Lee, Y. S. Beta and gamma binaural beats enhance auditory sentence comprehension. *Psychol. Res.* **87**, 2218–2227 (2023).
31. Wells, J., Christiansen, M., Race, D., Acheson, D. & Macdonald, M. Experience and sentence processing: Statistical learning and relative clause comprehension. *Cognit. Psychol.* **58**, 250–271 (2009).
32. Grahn, J. A. & Brett, M. Impairment of beat-based rhythm discrimination in Parkinson's disease. *Cortex* **45**, 54–61 (2009).
33. King, J. & Just, M. A. Individual differences in syntactic processing: The role of working memory. *J. Mem. Lang.* **30**, 580–602 (1991).
34. Noonan, N. B., Redmond, S. M. & Archibald, L. M. D. Contributions of children's linguistic and working memory proficiencies to their judgments of grammaticality. *J. Speech Lang. Hear. Res.* **57**, 979–989 (2014).
35. Droit-Volet, S., Wearden, J. H. & Zélandi, P. S. Cognitive abilities required in time judgment depending on the temporal tasks used: A comparison of children and adults. *Q. J. Exp. Psychol.* **68**, 2216–2242 (2015).
36. Bailey, J. A. & Penhune, V. B. Rhythm synchronization performance and auditory working memory in early- and late-trained musicians. *Exp. Brain Res.* **204**, 91–101 (2010).
37. Kotz, S. A., Schwartze, M. & Schmidt-Kassow, M. Non-motor basal ganglia functions: A review and proposal for a model of sensory predictability in auditory language perception. *Cortex* **45**, 982–990 (2009).
38. Buhusi, C. V. & Meck, W. H. What makes us tick? Functional and neural mechanisms of interval timing. *Nat. Rev. Neurosci.* **6**, 755–765 (2005).
39. Matell, M. S. & Meck, W. H. Cortico-striatal circuits and interval timing: Coincidence detection of oscillatory processes. *Cognit. Brain Res.* **21**, 139–170 (2004).
40. Iversen, J. R. & Balasubramaniam, R. Synchronization and temporal processing. *Curr. Opin. Behav. Sci.* **8**, 175–180 (2016).
41. Arnal, L. H. Predicting, “When” using the motor system's beta-band oscillations. *Front. Hum. Neurosci.* **6**, 225 (2012).
42. Van Petten, C. & Luka, B. J. Prediction during language comprehension: Benefits, costs, and ERP components. *Int. J. Psychophysiol.* **83**, 176–190 (2012).
43. Friederici, A. D. Event-related brain potential studies in language. *Curr. Neurol. Neurosci. Rep.* **4**, 466–470 (2004).
44. Goswami, U. A temporal sampling framework for developmental dyslexia. *Trends Cognit. Sci.* **15**, 3–10 (2011).
45. Tierney, A. & Kraus, N. Auditory-motor entrainment and phonological skills: Precise auditory timing hypothesis (PATH). *Front. Hum. Neurosci.* **8**, 949 (2014).
46. Goffman, L. Kinematic differentiation of prosodic categories in normal and disordered language development. *J. Speech Lang. Hear. Res.* **47**, 1088–1102 (2004).
47. Ladányi, E., Persici, V., Fiveash, A., Tillmann, B. & Gordon, R. L. Is atypical rhythm a risk factor for developmental speech and language disorders?. *Wiley Interdiscip. Rev. Cognit. Sci.* **11**, e1528 (2020).
48. Fiveash, A., Bedoin, N., Gordon, R. L. & Tillmann, B. Processing rhythm in speech and music: Shared mechanisms and implications for developmental speech and language disorders. *Neuropsychology* **35**, 771–791 (2021).
49. Macmillan, N. A. & Kaplan, H. L. Detection theory analysis of group data: Estimating sensitivity from average hit and false-alarm rates. *Psychol. Bull.* **98**, 15 (1985).
50. Bates, D., Mächler, M., Bolker, B. & Walker, S. Fitting linear mixed-effects models using lme4. *J. Stat. Softw.* (2015).
51. Fox, J. & Weisberg, S. *An R Companion to Applied Regression* (Sage Publications, 2018).

## Acknowledgements

This research was supported in part by the Neuroscience Innovation Foundation. We thank Sung-Joo Lim, Martin Wiener, Daeyeol Lee, and Kathryn V. Kreidler for helpful comments on the manuscript.

## Author contributions

H.W.K. and Y.S.L. designed research; H.W.K., J.K., J.S.B., and Y.M. performed research; H.W.K., M.M.M., and A.A. analyzed data; H.W.K. and Y.S.L. wrote the paper. All authors revised and edited the paper.

## Competing interests

The authors declare no competing interests.

### Additional information

**Correspondence** and requests for materials should be addressed to Y.S.L.

**Reprints and permissions information** is available at [www.nature.com/reprints](http://www.nature.com/reprints).

**Publisher's note** Springer Nature remains neutral with regard to jurisdictional claims in published maps and institutional affiliations.



**Open Access** This article is licensed under a Creative Commons Attribution 4.0 International License, which permits use, sharing, adaptation, distribution and reproduction in any medium or format, as long as you give appropriate credit to the original author(s) and the source, provide a link to the Creative Commons licence, and indicate if changes were made. The images or other third party material in this article are included in the article's Creative Commons licence, unless indicated otherwise in a credit line to the material. If material is not included in the article's Creative Commons licence and your intended use is not permitted by statutory regulation or exceeds the permitted use, you will need to obtain permission directly from the copyright holder. To view a copy of this licence, visit <http://creativecommons.org/licenses/by/4.0/>.

© The Author(s) 2024