Tapping to a beat in synchrony predicts brain print sensitivity in pre-readers

Paula Ríos-López[1], Nicola Molinaro[12] & Marie Lallier[1]

[1]BCBL, Basque Center on Cognition, Brain and Language, Paseo Mikeletegi, 69, 20009, Donostia-San Sebastián, Spain

[2]Ikerbasque, Basque foundation for Science, Bilbao, Spain

Corresponding author: Paula Ríos-López. Email address: p.rios@bcbl.eu

Address: Paseo Mikeletegi, 69, 20009, Donostia-San Sebastián, Spain

Abstract

This longitudinal study was aimed at testing the relation between rhythm sensitivity and behavioural and neural orthographic sensitivity in pre-reading stages. Basque-speaking children performed several behavioural and EEG tasks at two time points prior to formal reading acquisition (T1: 4 years old; T2: 5 years old). Neural sensitivity to print was measured via a novel child friendly N170-elicitation paradigm. Our results highlight a transversal and longitudinal relation between rhythm sensitivity and letter name knowledge in pre-reading children. Moreover, they show that children's rhythm sensitivity predicts a significant part of the variance of their N170 response one year later, highlighting the potential of rhythm tasks to predict future orthographic sensitivity in pre-reading stages. Interestingly, the relation between rhythmic skills and print sensitivity was not mediated by the children's phonological short-term memory. Our results provide novel evidence on the importance of rhythm sensitivity for the development of early orthographic sensitivity.

Keywords

Reading foundations, rhythm sensitivity, orthographic sensitivity, N170, cognitive development

1. Introduction

Reading is an instruction-derived ability that requires learning the cross-modal arbitrary mapping between speech sounds (phonemes) and their graphical representation. Although joint efforts from the educational and the scientific fields have been abundant, from 3 to 10% of children still struggle to acquire reading despite appropriate instruction and in the absence of further severe cognitive deficits (developmental dyslexia; Bishop & Snowling, 2004; Snowling, 2001). This being the case, an intense line of research has been devoted to the search of the *pre-requisites* of reading, i.e. the foundations upon which reading acquisition can consolidate. Within the field of pre-reading speech perception abilities, accumulating evidence points to a causal role of the slow speech rhythms perception for reading development (see Wood, Wade-Woolley, & Holliman, 2009 for a review). This idea is supported by abundant evidence in the field of dyslexia showing that, as compared to typically developing participants, dyslexic participants show deficits in behavioural receptive and productive rhythm tasks (Corriveau & Goswami, 2009; Corriveau, Goswami, & Thomson, 2010; Flaugnacco et al., 2014; Huss, Verney, Fosker, Mead, & Goswami, 2011; Richardson, Thomson, Scott, & Goswami, 2004; Thomson, Fryer, Maltby, & Goswami, 2006; Thomson & Goswami, 2008; Tierney & Kraus, 2013). Despite the importance of the evidence coming from the comparison between dyslexic and typically developing participants, it does not allow discarding the presence of bidirectional contributions between reading acquisition and rhythm sensitivity. More appropriate evidence on the causal relation between these two comes from longitudinal studies that established a relation between rhythm sensitivity prior to reading acquisition and future reading, and from transversal studies reporting a tight relation between rhythm sensitivity and other classical reading predictors in pre-readers (e.g. Holliman et al., 2010; Lundetræ & Thomson, 2017; Moritz et al., 2013; Ozernov-Palchik, Wolf, & Patel, 2018; Carr, White-Schwoch, Tierney, Strait, & Kraus, 2014).

In this context, several theories have proposed that the relation between rhythm sensitivity and reading predictors in pre-reading stages would be mediated by phonological skills. Specifically, they suggest that the sensitivity to language rhythm would aid speech segmentation (Ramus, Hauser, Miller, Morris, & Mehler, 2000) bootstrapping hence phonological awareness, which would in turn boost reading development (Goswami, Power, Lallier, & Facoetti, 2014; Lallier et al., 2018; Lallier, Molinaro, Lizarazu, Bourguignon, & Carreiras, 2017; Leong, Hämäläinen, Soltész, & Goswami, 2011; Moritz et al., 2013; Wood & Terrell, 1998). Following this theoretical line, studies examining the relation between rhythm sensitivity and other pre-reading skills have focused predominantly on phonological reading predictors, i.e. on phonological awareness, phonological short-term memory and RAN and have indeed found a tight relation between these and rhythm sensitivity in pre-readers (e.g. Carr et al., 2014; Holliman et al. 2010; Moritz et al., 2013).

The present study will address the issue as to whether rhythmic sensitivity plays a role in the development of early (pre-reading) *orthographic predictors* of reading. This question is timely and of interest given recent evidence challenging the mediator role of phonological processing skills in the rhythm-reading relation (Calet, Gutierréz-Palma, Defior, & Jimenéz-Fernández, 2019; Calet, Gutiérrez-Palma, Simpson, González-Trujillo, & Defior, 2015; González-Trujillo, Defior, & Gutiérrez-Palma, 2014; Holliman et al., 2014; Holliman, Wood, & Sheehy, 2008). For example, Holliman et al. (2008) reported that linguistic rhythm sensitivity (as assessed by a stress manipulation task) predicted significant variance at early stages of reading development after controlling for performance in a phonemic awareness task. The authors proposed that early segmental phonological skills might be relatively independent of early orthographic skills such that very young children might develop (rudimentary) pre-reading phonological awareness via exposure to speech since birth, whereas orthographic representations of language might develop independently as a function of experience with written language. Accordingly, Wood (2004) reported no differences in phonological awareness between groups of pre-reading children classified as a function of their knowledge of the letters of the alphabet. Without denying the relation between phonological awareness and orthographic skills once reading is acquired, this evidence opens the possibility that during pre-reading stages, the contribution of rhythm sensitivity to orthographic sensitivity could be partly independent of phonological skills. In fact, rhythmic sensitivity might contribute to early reading and orthographic development through two pathways amongst which only one would be mediated by phonology (Holliman et al., 2008). Although identifying potential non-phonological mediators falls out of the scope of the present study, some have suggested that factors such as written prosody (e.g., Calet et al., 2019) or morphology (Holliman et al., 2014) could be good candidates.

In pre-readers, one study to date has specifically assessed the relation between rhythmic abilities and orthographic predictors of reading such as letter knowledge (Ozernov-Palchik et al., 2018). This study showed that pre-readers' (mean age: 5.8 years) performance on a metrical perception task accounted for unique variance in a letter-sound-knowledge task, and perhaps more interestingly, even beyond phonological awareness. In any case, the authors tested the children only at one time point, which prevents from concluding on the causality of this relation. The debate on the independent contribution of rhythm sensitivity to orthographic sensitivity at pre-reading and early reading stages is hence open. One way to examine this question is to test both rhythm and orthographic sensitivity when the latter is assumed to be relatively independent of phonemic awareness, i.e. before formal reading tuition (Wood, 2004). To our knowledge, no study has taken such enterprise from a longitudinal perspective to establish a causal relationship between rhythmic processing and print sensitivity, controlling for phonological processing skills.

Beside behavioural measures of orthographic sensitivity such as letter knowledge (Byrne & Fielding-Barnsley, 1990; Evans, Bell, Shaw, Moretti, & Page, 2006; Stuart & Coltheart, 1988), pre-readers' orthographic sensitivity can also be measured at the brain level, whilst performing orthographic tasks under electrophysiological recording. Measuring behavioural performance in very young children is often prone to high variability that does not only result from the process of interest but that is likely to stem from variability in general speed of motor/oral response or strategic response bias. Event related potentials (ERP) are generally more sensitive to capture variability in certain processes since they are able to measure very early effects elicited by the unconscious processing of the stimuli free from any response bias. In the present study, we assessed the contribution of rhythmic sensitivity to both a behavioural, i.e., letter name knowledge (LNK) and a neural, i.e., the N170, early orthographic predictors or reading.

The N170 is a negative ERP that appears around 170 milliseconds after the presentation of print as opposed to other control stimuli such as symbol strings or false fonts (Brem et al., 2009; Maurer, Brandeis, et al., 2005; Maurer, Brem, et al., 2005; Maurer et al., 2006). This ERP component has been repeatedly shown as one neurophysiological correlate of expert (Bentin, Mouchetant-Rostaing, Giard, Echallier, & Pernier, 1999; Brem et al., 2005; Maurer, Brandeis, & McCandliss, 2005) and developing (Bach, Richardson, Brandeis, Martin, & Brem, 2013; Brem et al., 2013; Lochy, Van Reybroeck, & Rossion, 2016; Maurer et al., 2006) print sensitivity. Although it had been proposed that the N170 response should be absent in pre-readers (Brem et al., 2009, 2005; Eberhard-Moscicka et al., 2015; Maurer, Brem, et al., 2005; Maurer et al., 2006), several studies have challenged this assumption (Bach et al., 2013; Brem et al., 2013; Lochy et al., 2016). For example, Lochy et al. (2016) reported that letters embedded in pseudofont streams elicited a pronounced response in left occipital electrodes in pre-readers and that individual differences in the amplitude of the letter-specific response was positively correlated with children's LNK. Additionally, pre-reading N170 amplitude difference between words and symbols were found to correlate positively with children's LNK (Maurer, Brem, et al., 2005). Lastly, Brem et al. (2013) found that after a short training in letter-to-sound correspondences, pre-readers showed a significant N170 to words in contrast to symbols in the left hemisphere.

Overall, the aforementioned evidence suggests that print-specific N170 as well as LNK could be two good indexes of orthographic sensitivity in the pre-reading brain and are good candidates to test the contribution of rhythmic sensitivity to pre-reading orthographic sensitivity.

This study

The goal of this longitudinal study was to test the relation between rhythm sensitivity and orthographic sensitivity at pre-reading stages, whilst controlling for the putative mediation of phonological skills in the relation. On the one hand, we wanted to examine the relation between rhythm sensitivity and a behavioural orthographic marker of reading readiness (i.e., LNK), for which evidence is scarce in cross-sectional designs (Ozernov-Palchik et al., 2018) and absent in the longitudinal domain. On the other hand, we intended to extend the relation between rhythmic abilities and orthographic reading readiness at the neural level (i.e., N170 responses).

In that aim, we carried out a longitudinal study in which we collected pre-readers data at two different times before children received formal reading instruction: during the summer breaks immediately previous (T1; four years old) and posterior (T2 five years old) to the last year of kindergarten. Both at T1 and T2, we collected children's performance on a LNK task and on an online beat synchronization task previously used in the literature to measure rhythm sensitivity (Carr et al., 2014). In addition, we collected two phonological measures (Phoneme awareness and Phonological short-term memory) to test the hypothesis that the rhythm-reading readiness relation should be mediated by phonological skills since pre-reading stages (e.g. Goswami, 2011). At T2, we also collected the electrophysiological print-specific responses of the children using EEG. Rather than a one-back task with whole words, we used a passive viewing paradigm with consonant-vowel syllables. Indeed, some (Lochy et al., 2016) have proposed that the presence of the N170 in pre-readers could be obscured by the high perceptual and memory demands of classic one-back tasks using whole words (e.g. Maurer, Brem, et al., 2005; Maurer et al., 2006).

Hypotheses

Firstly, if rhythmic sensitivity plays a role in pre-reading behavioural orthographic sensitivity, we expected performance on the beat synchronization task to be correlated with LNK before the formal start of reading tuition, both within T1, within T2, and across T1 and T2 (Ozernov-Palchik et al., 2018).

Secondly, if rhythmic sensitivity plays a role in pre-reading neural orthographic sensitivity, and if N170 responses are a good neural index of print sensitivity at the pre-reading stage, we expected beat synchronization at T1 to explain a significant part of the variance in the amplitude of the N170 at T2, as well as a significant relation between beat synchronization and N170 at T2.

Thirdly, if LNK and N170 both reflect a unique orthographic sensitivity construct, we expected significant relations between LNK measured at T1 and T2 with the N170 responses measured at T2.

Lastly, we had no clear prediction as to whether we would find a significant mediating effect of phonological skills in the relation between beat synchronization performance and both our behavioural (LNK) and neural (N170) markers of orthographic sensitivity. On the one hand, several studies found a relation between rhythmic abilities and phonological processing (e.g. Carr et al., 2014; Holliman et al. 2010; Moritz et al., 2013). On the other hand, recent evidence challenged the idea that phonological skills could mediate the relation between rhythm sensitivity and early reading and orthographic development (Calet et al., 2015; Holliman et al., 2008; Ozernov-Palchik et al., 2018).

2. Materials and methods

2.1. Participants

Forty-three pre-readers (22 males; age *M* = 5.09 years; *SD* = 0.31) were recruited from two different schools in Donostia-San Sebastián (Basque Country, Spain).

Thirty-eight children came back for behavioural testing one year later (20 males; age *M* = 5.88 years; *SD* = 0.34). Thirty-two of these children completed also the EEG session.

The Basque Country is a bilingual community in which two official languages coexist (Basque and Spanish), but given that all children received formal education in Basque, all linguistic tasks were performed in this language.

Testing proceeded during the summer vacations prior to (T1) and after (T2) the last year of kindergarten. At T1, children completed only the behavioural tasks. At T2, children came to the lab twice; the first time for the behavioural session and the second time for the EEG session.

At both testing times, parents were asked to fulfil a questionnaire on reading habits to ensure that children could not read fluently in any of the phases of the longitudinal study. Moreover, we asked children directly whether they knew how to read and ran a quick test consisting on the presentation of short words printed in paper. We considered that failing to complete this test was a sign of the child's inability to decode written material. The explicit question and the short test confirmed that none of the children that participated in this study was able to read.

All children were tested under signed parental authorization and the experiment was approved by the BCBL Ethics Review Board and complied with the guidelines of the Helsinki Declaration.

2.2. Stimuli and procedure

2.2.1. Behavioural battery (T1 and T2)

Children completed five tasks while seating comfortably in a shielded room with the experimenter. The session took approximately 30 minutes.

2.2.1.1. *IQ assessment*

The Matrix reasoning subtest from the WPPSI-III (Wechsler, 2002) was used to obtain a measure of non-verbal IQ. Children were presented with a matrix of three drawings and a blank square and were asked to complete the series with the correct element out of four to five alternatives. The test consisted of a total of 29 items, and the initial item was adapted upon the child's age. Direct scores where converted into scalar (standardized) scores to correct for children's age differences (in months) and for comparability across testing years.

2.2.1.2. Reading predictors

i) Letter name knowledge (LNK)

The letters of the Basque alphabet were presented in capital case and one by one in a random order. The child’s task was to name the letter. The experimenter recorded accuracy scores (maximum of 27).

ii) Phoneme deletion/elision

This task has been extensively used to measure phonological awareness in young children (e.g. Stanovich, Cunningham, & Cramer, 1984). Stimuli consisted of a total of 16 pseudowords and were presented one at a time through headphones at 80 dB SPL. Immediately after presentation, children were asked to recall them, but omitting the first phoneme of the pseudoword (phoneme deletion) for half of the items (eight) and the first phoneme of the pseudoword’s second syllable (phoneme elision) for the remaining half. For each subtask, stimuli consisted of eight by-syllabic items that were presented in a randomised order. Within the deletion task, half of the items (four) consisted of two syllables with a consonant-vowel structure (CV; e.g. /timu/; correct response: /imu/). The remaining half of the items had a consonant-consonant-vowel (CCV) structure in the first syllable and a simple CV structure in the second (e.g. /bluno/, correct response: /luno/). The structure of the stimuli for the elision subtask was identical to the deletion subtask (i.e. four CV-CV items and four CCV-CV items). Children now had to omit the first phoneme of the second syllable (e.g. /bupa/; correct response: /bua/) and consequently, they never had to *break down* a consonant cluster (i.e. /flope/; correct response: /floe/). This was done in order to decrease the task complexity due to our participants’ young age. The number of correct responses was recorded (maximum of 16).

iii) Pseudoword repetition

This task has been extensively used to measure phonological short-term memory and its contribution to reading skills (Jorm, 1979; Muter & Snowling, 1998). Stimuli were presented through headphones at an SPL of 80 dB. Stimuli consisted in a total of 28 items that complied with Basque phonotactics and the child was instructed to repeat the corresponding item immediately after he/she heard it. Stimuli consisted in items differing in length: 8 two-syllable items, 8 three-syllable items, 6 four-syllable items, and 6 five-syllable items. Seventy per cent of the stimuli were formed by syllables whose phonemic structure followed the consonant-vowel pattern. For the thirty per cent of remaining stimuli, one syllable contained a consonant cluster (three letters; e.g. /pra/). Items were different in T1 and T2 to avoid (not likely) repetition effects across the years. In each testing time, two practice items were presented before the task started to ensure that the children understood the task. This task took approx. five minutes. Total time for the repetition of the whole list and errors were recorded by the experimenter. A measure of efficiency per minute was calculated as follows:

$$\text{Pseudoword efficiency} = \frac{\text{Total number of correct pseudowords}}{\text{Total time}} * 60$$

iv) Beat synchronisation

To assess the rhythmic skills of the children, we adapted a task that has been previously shown to correlate with other reading predictors in pre-readers (Carr et al., 2014). The task was originally proposed by Kirschner and Tomasello (2009) as the best available task to measure children's motor skills in social situations. In this task, the child is asked to tap to a beat in synchrony with the experimenter. This situation, due to the social interaction with another human being, generates a special motivation for synchronised movement that is absent when tapping to a recorded rhythm (Kirschner & Tomasello, 2009). In our case, the experimenter produced a rhythmic drum beat and children were asked to "synchronise" to this rhythm, i.e. to tap at the tempo set by the experimenter. Children were asked to tap with their dominant hand, since significant group differences when tapping to a beat with the dominant *versus* the non-dominant hand have been previously reported in the literature (Nozaradan, Zerouali, Peretz, & Mouraux, 2015). Before the task, the experimenter was trained to produce the rhythmic pattern as accurately as possible, so that the mean *SD* of her inter-drum intervals was below 25 ms. The task consisted of two trials at 1.67 Hz (606 ms ISI; duration of 20 seconds; total 50 beats). The experimenter and the child produced the drum beats separately in a peripheral drumming pad. Input was recorded via a custom-script created in *Python*.

2.2.2. EEG session (T2)

2.2.2.1. Syllable presentation task

This task was used to record electrophysiological activity evoked by print (syllables) as compared to other perceptually similar but non-meaningful visual stimuli (false fonts) before formal reading instruction. Stimuli consisted of 72 syllables and 72 false fonts created via Adobe Illustrator with font Consolas Regular and size 200. Syllables were formed by a combination of consonants /b/, /p/, /d/, /t/, /g/ and /k/ with vowels /a/, /e/, /i/, /o/ and /u/. 12 of the syllables (*ba, be, di, do, ga, gu, pe, pi, to, tu, ka, ke)* were presented three times and 18 of the syllables (*bi, bo, bu, da, de, du, ge, gi, go, pa, po, pu, ta, te, ti, ki, ko, ku*) two times to make a total of 72. False fonts were constructed rotating and inverting the features of original syllables to avoid low level visual differences between experimental conditions (see Supplementary Figure 1 for a set of examples of the stimuli presented in each condition). Each false font was repeated three times. Fillers (n = 20) consisted of the image of a dog and were included to ensure that children were paying constant attention to the screen. Presentation order was pseudorandomised, such that more than three stimuli of the same time did not appear consecutively, and delivered through *Psychopy* (Peirce, 2009). The children's task was to press the "space" key every time the dog appeared in the screen. All types of stimuli appeared in the screen for 1000 ms and interstimulus interval was set at 0 ms. Ten practice trials were included before starting with the experiment to ensure that the child understood the task. The total duration of the task was 4 minutes.

2.2.2.2. EEG data acquisition

The EEG signals were recorded in a child-friendly soundproof electrically-shielded room. Children were comfortably seated in a chair adapted to their size. They were instructed to stay silent and to avoid moving as much as possible. EEG signals were recorded using a Brain Products GmbH actiCAP with 32 electrodes. To reduce preparation time and avoid children's fatigue, we reduced the array of scalp electrodes to 19 (FP1, FP2, F7, F3, Fz, F8, F4, C3, Cz, C4, T7, T8, P7, P3, Pz, P8, P4, O1 and O2) which were distributed over the scalp based on 10-10 International System. Additional reference electrodes were placed on both mastoids A1 and A2. In addition, we recorded electrode FCz and around the eyes at left Heog, right Heog and left Veog. Signals were sampled at 500 Hz with a high-pass filter at 0.1 Hz. Monopolar differential recording was referenced online to electrode FCz.

2.3. Data analysis

2.3.1. Behavioural data

We considered that children were outliers in a task when their score was above 3 *SD* from the mean of the group. Furthermore, we also assessed the children's performance across testing times to spot longitudinal outliers (i.e. children whose performance from T1 to T2 decreased inexplicably) using the same criterion (≥ 3 *SD*).

2.3.1.1. Beat synchronisation task: Drumming consistency

Data analysis of the BSynch task was based on circular statistics (Fisher, 1993). Circular statistics are intended to assess the level of synchronization to the pace of an external beat (see Kirchner & Tomasello, 2009 for a full justification for the use of this type of statistics with children data). We recorded both onset and offset movements, but for the purpose of this study only onsets were analysed. Data analysis consisted of the following steps: i) Drum hits were assigned a point in a circular scale, with the stimulus beat aligned at 0º. Every beat was defined by an angle in degrees or phase, i.e. the result of subtracting the onset time from the onset of the stimulus closest in time. In this way, it is possible to measure the degree of synchronization independently of its attribution to the preceding or following stimulus drum hit, i.e. instead of using traditional linear statistics, a measure of phase consistency is created; ii) thereafter, the result was divided by the ISI and multiplied by 360; iii) lastly, all vectors were summed and divided by the number of hits generated, which rendered a mean vector *R*. The mean vector *R* gives information about two parameters: the vector's mean direction *h*, which can be used as a measure of the child's phase preference, and the vector's mean length. The latter provides a measure of the phase consistency of the produced drum hits with regard to the stimulus. Its value ranges between 0 (no phase consistency) and 1 (perfect phase consistency). For the purpose of this experiment, we only analysed the vector *R*'s mean length, which was *z*-transformed using the function *circ_rtest*() available in Matlab's package *CircStat* (Berens, 2009).

2.3.1.2. Joint analysis of reading predictors

The aim of this part of the analysis was to explore concurrent and longitudinal correlations among our behavioural tasks. First, scores were *z*-transformed for comparability among tasks. Partial correlations controlling for IQ and age were performed to examine the test-retest reliability of the Beat synchronisation and the LNK tasks and to explore the inter-correlation between these variables. When the specific analysis involved multiple comparisons, both uncorrected and FDR-corrected *p* values are reported. Lastly, based on previous theoretical frameworks claiming that the relation between early literacy indexes and rhythm sensitivity might be mediated by phonological skills (Goswami, 2011), we conducted a second analysis in which the children's phonological skills (phonological short-term memory, see Results section) were controlled for.

2.3.2. EEG data

We used BrainVision Analyzer 2.0 to pre-process the EEG data. Statistical analyses of the EEG data were carried out in R (R Core Team, 2017). Based on previous studies on the N170 component, we focused our analysis on a broad time interval between 100 and 400 ms (e.g. Bach et al., 2013; Maurer, Brandeis, et al., 2005). Due to the facts that the N170 effect had never tested before in Basque-speaking pre-readers and that we used a novel task, we run an objective statistical test to refine our time window of interest to the interval where the differential N170 to print was actually present. The test consisted in computing multiple point-by-point t-tests (FDR-corrected for multiple comparisons) to compare the amplitude evoked by Syllables and by False fonts across the whole time window (from -200 to 400 ms). Next, we averaged the data across our window of interest and used the *lme4* package to build a mixed effect model (Baayen, Davidson, & Bates, 2008) with the amplitude of the EEG as dependent variable. Concerning the construction of the random effects, the complexity of their structure was increased in a series of models (Barr, Levy, Scheepers, & Tily, 2013), leading to a maximal model that converged and contained only the random effects that increased significantly the residual variance captured by the model. The fixed effects structure was built up with the Condition (Syllable vs. False font) and the Hemisphere (Left vs. Right) factors and their interaction. In addition, we always introduced to the model the factors IQ (continuous Matrix score) and Age (in months) to control for intelligence and age effects in the amplitude of the EEG signal. Regarding post-hoc analysis for exploring interaction effects, we set up paired contrasts using the least squares means method available in the R package *lsmeans* (Lenth, 2016). This method allows estimating pairwise comparisons of the slopes of categorical variables fitted in a linear mixed effect model at different levels of a continuous variable. In this analysis, significance estimation is based on confidence intervals at 0.95 and corrected for multiplicity with the Tukey HSD method. Finally, as an indicator of the models' goodness of fit, we estimated the coefficient of determination (*pseudo* R^2) and report both the marginal R^2 (i.e. the proportion of variance explained only by the fixed effects) and the conditional R^2 (i.e. the proportion of variance explained by both the fixed and the random effects together) for each specific fit.

After inspection of the raw data, we decided to discard directly the data of ten participants due to excessive movement during recording or to reduced number of trials. Therefore, the total number of children whose data were considered for further analysis was 22. Further raw inspection of the data indicated the presence of excessive noise in mastoid channels due to movement. Accordingly, we excluded electrodes A1 and A2 from further analysis, and re-referenced the data offline to electrode Cz. In each individual case, we pooled other channels with excessive noise from the signal of at least three surrounding electrodes. Next, we applied the signal a Butterworth Zero Phase Bandpass Filter (0.1 – 45 Hz; 48 dB/oct) and a notch filter at 50 Hz. Then, we inspected manually the raw data to detect ocular activity and head movements. Afterwards, we applied an Independent Component Analysis (ICA) analysis (Infomax; Bell & Sejnowski, 1995) to the whole participant's recording. We then inspected the components to detect blinks, heart beat and muscular activity and removed them accordingly. We defined the epochs of interest from 200 ms before to 400 ms after stimulus presentation and proceeded with semiautomatic artefact rejection. We removed segments containing voltage steps larger than 50 µV/ms and/or voltage differences of more than 150 µV in intervals of 200 ms and segments where activity was lower than 0.5 µV for 100 ms or more. These criteria led us to discard a total of 21% of the epochs (with no difference between the Syllables and the False font conditions in the amount of rejected trials; $t(21) = 0.14$; $p = 0.89$). At this point, we discarded individual data sets for which more than 25% of the data for one or both conditions were rejected (i.e. lower threshold for further analysis was set at 75% of the data, or trials 54 trials, per condition). We discarded the data of three children based on this threshold. Thereafter, we baseline corrected (200 ms to 0 ms) the data of the 19 children for whom enough movement-free trials were available. We calculated individual ERPs for each condition and averaged them across subjects to calculate grand-averaged ERPs.

2.3.2. Behavioural and EEG data

After the initial analyses of the separate EEG data, we moved on to building mixed effects models that included our behavioural tasks of interest. In all analysis, we included Beat synchronisation (our main predictor of interest) and LNK (due to the relation between this and the N170 found previously in the literature; e.g. Lochy et al., 2016).

Models were built separately for the different testing times (i.e. one model including the behavioural predictors at T1 and a second model including the behavioural predictors at T2). The dependent variable of the models was the amplitude of the N170 effect at T2 averaged across the time window of interest, and the fixed effects were built by the factors Condition (Syllable vs. False font), Hemisphere (Left vs. Right), continuous score in the Beat synchronisation task (at T1 or T2, correspondingly) and continuous score in the LNK task (at T1 or T2, correspondingly), and all possible interactions between these factors. As for the previous analyses, IQ and age were also introduced in the model to control for effects of these factors on the amplitude of the EEG signal. The procedure followed to explore post-hoc contrasts was identical to the one followed for the EEG-only analysis.

Lastly, we built two models (one per testing time) in which, beside our behavioural predictors of interest (LNK and Beat synchronisation), the score of the children in the Pseudoword repetition task was introduced. Models with and without the phonological factor were compared through analyses of variance (ANOVA) to test if introducing this factor improved significantly the model’s fit of the data, suggesting that the phonological skills explained part of the N170 variance beyond our behavioural predictors of interest.

3. Results

Within the Results section, we first present the descriptive and correlational analysis of the behavioural data (section 3.1). Section 3.2 presents the separate analysis of the Syllable presentation task (EEG data). Finally, section 3.3 introduces different models in which the behavioural data at the different testing times are used to predict the variance of the EEG data.

3.1. Behavioural data

3.1.1. Descriptive statistics

Unfortunately, the Phoneme deletion/elision task appeared to be too difficult for the children and they were not able to complete it successfully in any of the testing times. Therefore, the score in the phonological short-term memory task (Pseudoword repetition) was the only phonology-related measure available for analysis (see Discussion).

The longitudinal performance of five of the children in the Beat synchronisation task was notably deviated from the mean of the group from the longitudinal perspective, such that there was an inexplicable detriment (≥ 3 *SD*) in performance from T1 to T2. In consequence, we considered that the risk of measurement error or experimental fatigue in these five participants was high and did not use their data for further analysis. Table 1 shows the scores of the children in the behavioural tasks after outlier removal.

Table 1. Behavioural results after outlier removal

	T	n	*M*	*SD*	median	min	max	kurtosis
Matrices	1	38	7.74	2.67	8	3	14	-0.40
	2	31	6.39	2.36	6	2	11	-0.92
LNK	1	36	14.86	6.13	16	1	23	-1.07
	2	31	20.26	3.43	21	7	23	4.78
BSynch*	1	34	1.69	1.38	1.63	0	4.87	-0.32
	2	31	4.88	5.84	2.22	0.25	21.57	1.80

*Z-transformed as specified in text (*2.3.1.1. Beat synchronisation task: Drumming consistency*); BSynch = Beat synchronisation

3.1.2. Partial correlations

As shown in Table 2, our different measures correlated significantly with themselves across testing times, showing hence reliability. Table 2 shows the results of the partial correlations between our different measures within and across testing times. (see Supplementary Figure 2 for a visual depiction of all correlations presented in this section).

Table 2. Partial correlations between the LNK and the Beat synchronisation tasks across testing times

	1.			2.			3.		
	n	***r***	***p* (FDR)**	**n**	***r***	***p* (FDR)**	**n**	***r***	***p* (FDR)**
1. LNK T1			-						
2. LNK T2	30	.60	<.001*			-			
3. BSynch T1	33	.26	.14	27	.36	.06 (.08)*			-
4. BSynch T2	29	.50	<.01 (.02)*	27	.49	.01 (.02)*	29	.46	.01*

BSynch = Beat synchronisation. FDR correction (in brackets) is only provided when relevant (i.e. when multiple comparisons are involved).

3.1.3. Phonological short-term memory as a mediator in the relation between rhythmic skills and letter knowledge

After controlling for children's score in the Pseudoword repetition task, the partial correlation coefficients between LNK and Beat synchronisation did not change significantly either at T1 (R = .30; p = .15) or at T2 (R = .47; p = .02), suggesting that, in our sample, the relation between Beat synchronisation and LNK was not mediated by the scores of the children in this phonological task.

3.2. EEG data

Visual inspection of the topographical maps across the time window ranging from -200 ms (baseline) up to 400 ms after stimulus presentation showed that the maximum difference of amplitude between our conditions of interest (Syllables vs. False fonts) occurred in the window between 190 and 325 ms in left posterior (P7, P3 and O1) and right posterior (P8, P4 and O2) electrodes (see Figure 1).

[Insert Figure 1 around here]

After visual inspection, we transferred the data to R (R Core Team, 2017) for statistical analysis. In the interest of reducing the number of statistical comparisons, we computed the mean of the left and right electrodes separately and created two scalp regions: the left hemisphere region (LH region; electrodes P3, P7 and O1), and the right hemisphere region (RH region; electrodes P4, P8 and O2). The ERP analysis for the separate electrodes can be consulted in the Supplementary analyses in the Appendix.

First, to support our selection of the N170-time window with an objective measure and to ensure that no other differences between the waves was present before the N170 component, we computed a point-by-point (in steps of 2 ms; sampling frequency = 500 Hz) paired *t*-tests across the whole window (-200 to 400 ms) in both regions (see Figure 1). Results showed that in the window between -200 ms and 190 ms (i.e. were no significant difference was expected), there was no difference between the waves at any time point in any of the regions (all *p*s > 0.05). The significant difference between Syllables and False fonts started earlier in the LH region (at 198 ms; $t(18) = 2.70$; FDR corrected $p < 0.05$) as compared to the RH region (at 222 ms; $t(18) = 2.58$; all $p < 0.05$). Regardless of region, the difference between the waves was sustained until the end of our window of interest (325 ms; all *p*s < 0.05). Finally, visual inspection of the waves indicated that at around 325 ms, although the experimental effect between conditions was still present, ERPs showed a subsequent positive trend reflecting a later positive component following the N170. Differences in this later time interval could reflect a carry-over of the differences observed in the N170 interval. Accordingly, the differences across waves was significant in all electrodes at all time points between 325 and 400 ms when the baseline was kept at -200 ms (all *p*s < 0.05), but it did not reach statistical significance when the baseline was set at 325 ms (all *ps* < 0.05). Hence, this analysis confirmed that the differences after 325 ms were most probably a post-effect of the previous N170 component, reason why we did not consider for further analysis any activity occurring after this time point. For later analyses, we selected our time window of interest to include the whole range of significant time points found across the regions, i.e. from 198 to 325 ms.

Thereafter, we moved on to explore the possibility that the activity prompted by Syllables and False fonts was different as a function of region (Hemisphere). To this aim, we fitted a linear mixed effect model (Baayen et al., 2008) with average amplitude of the EEG signal in our window of interest as dependent variable, and Condition (two levels: Syllables and False fonts), Hemisphere (two levels: Left and Right) and their interaction (Condition by Hemisphere) as main predictors. Reference levels for these predictors were the False fonts for the Condition factor, and the Left hemisphere for the Hemisphere factor. The fixed effects structure also included factors IQ and Age. The maximal converging model included a random intercept by-Subject and two random slopes by-subject for the effect of the factors Condition and Hemisphere (to control for within-subject variability and measures' interdependency).

Table 3 shows the full summary of the model's results. The analysis showed that there was a significant main effect of Condition, such that the EEG amplitude values were more negative for Syllables (mean = -.48, *SD* = 6.20) as compared to False fonts (mean = 4.08, *SD* = 4.10), whereas we found no amplitude differences between hemispheres to the different stimuli.

Table 3. Summary of results of the model with amplitude of the EEG at T2 as DV and Condition, Hemisphere and their interaction as predictors

		Model characteristics			
	n	**obs**	**M R^2**	**C R^2**	
	19	76	.48	.93	
		Random effects			
		SD			
By-subject intercept		3.15			
By-subject slope Condition		3.63			
By-subject slope Hemisphere		1.96			
Residual		1.81			
		Fixed effects			
		β	**SE**	***t***	***p***
Intercept		-59.93	15.09	-3.97	<.01
Age		.89	.24	3.79	<.01
IQ		1.29	.30	4.32	<.001
Condition[1]		-5.34	1.10	-5.24	<.001
Hemisphere[2]		.84	.74	1.13	.26
Condition * Hemisphere		.75	.83	.90	.37

[1]Reference: False fonts; [2]Reference: Left hemisphere. C R^2 = Conditional R square; M R^2 = Marginal R square; obs = total of observations

3.3. Behavioural and EEG data

3.3.1. Longitudinal analysis

This analysis was carried out with the data of 17 children for whom data for all the tasks involved were available (to avoid missing data biases). Before fitting the model, we calculated the variance inflation factor (VIF) of our predictors to assess possible multicollinearity issues, given that our behavioural predictors (Beat synchronisation and LNK) were correlated. This analysis showed that our data were not severely affected by multicollinearity, since the VIF values of our variables were below the maximum recommended value of five (VIF of LNK: 4.56; VIF of Beat synchronisation: 4.44; Rogerson, 2001).

For parsimony reasons, only the relevant main fixed effects and interactions will be discussed in the text (see Table 4 a for a summary of the statistics of all the model's effects). All interactions involving Beat synchronisation at T1 were modulated by the significant triple interaction Beat synchronisation T1 by Condition by Hemisphere, such that response to Syllables as compared to False fonts was different in the Left and the Right hemispheres as a function of performance in the Beat synchronisation task. As shown in Table 4 a, neither the main effect of LNK at T1 nor any of the interactions in which it was involved were significant. In consequence, we focused our post-hoc analysis in the triple interaction Beat synchronisation T1 by Condition by Hemisphere (see Table 4 b and left panel of Figure 2). The Tukey HSD comparisons indicated that, in the Left hemisphere, the slopes for the Syllables and the False fonts were different for varying levels of performance in the Beat synchronisation such that the negativity for the Syllables increased as the score in the Beat synchronisation task increased, while this trend was not present for the False fonts (see Supplementary table 3 for description of the least square means for the different factors' levels). On the contrary, in the Right hemisphere, neither the slope for the Syllables nor the slope for the False fonts changed significantly with varying levels of the Beat synchronisation task.

Table 4

a. Summary of results of the model with amplitude of the EEG at T2 as DV and Condition, Hemisphere, Beat synchronisation T1, LNK T1, Age and IQ as predictors

Model characteristics				
	n	**obs**	**M R^2**	**C R^2**
	17	68	.55	.95
Random effects				
	SD			
By-subject intercept	3.00			
By-subject slope Condition	2.61			
By-subject slope Hemisphere	2.21			

Residual	1.44			
Fixed effects				
	β	**SE**	***t***	***p***
Intercept	-62.67	14.06	-4.46	<0.001*
BSynch T1	-.73	1.45	-50	0.62
Condition[1]	-6.20	0.87	-7.12	<0.001*
Hemisphere[2]	-1.35	0.79	-1.70	0.10
LNK T1	0.51	0.90	0.57	0.58
Age	0.97	0.22	4.49	<0.001*
IQ	1.08	0.32	3.35	<0.01*
BSynch T1 * Condition	-3.51	1.36	-2.58	0.02*
BSynch T1 * Hemisphere	-.53	1.24	-.43	0.67
Condition * Hemisphere	1.88	0.76	2.48	0.03*
LNK T1 * Condition	-1.37	0.81	-1.69	0.10
LNK T1 * Hemisphere	0.16	0.74	0.22	0.83
BSynch T1 * Condition * Hemisphere	2.88	1.19	2.43	0.03*
LNK T1 * Condition * Hemisphere	2.88	0.71	2.43	0.95

[1]Reference: False fonts; [2]Reference: Left hemisphere. BSynch = Beat synchronisation; C R^2 = Conditional R square; FF = False fonts; M R^2 = Marginal R square; obs = total of observations; Syll = Syllables

b. Tukey HSD post-hoc comparison of the triple interaction Beat synchronisation T1 by Condition by Hemisphere

Contrast	**β**	**SE**	***t***	***p***
	Left Hemisphere			
False fonts - Syllables	3.51	1.36	2.58	.02
	Right Hemisphere			
	.63	1.36	.46	.64

[Insert Figure 2 around here]

3.3.2. Within-T2 analysis

The procedure for model construction and analysis of the within-T2 data was identical to the one of the previous section, with the exception that the behavioural predictors (Beat synchronisation and LNK) were now the ones corresponding to T2. In this model, the triple interaction Beat synchronisation T2 by Condition by Hemisphere was again significant. As compared to the longitudinal model, LNK at T2 interacted significantly with the Condition factor. A description of all fixed effects results for this model can be seen in Table 5 a.

We then moved on to explore the significant interactions in the model. Regarding the triple interaction Beat synchronisation T2 by Condition by Hemisphere (i.e. the pairwise comparisons of the levels of Condition and Hemisphere as a function of varying scores in Beat synchronisation T2), the least square means analysis showed that, as for the longitudinal data, amplitude for the Syllables was more negative for higher scores at the Beat synchronisation task in the Left hemisphere, while this effect was absent for the False fonts (see Table 5 b and Figure 3). Nevertheless, this comparison did not survive the Tukey HSD adjustment for multiplicity (see Supplementary table 4 in the Appendix for description of the least square means for the different factors' levels). As shown in Figure 3, the amplitude for both the Syllables and the False fonts was more negative with higher performance in the Beat synchronisation task. Regarding the interaction LNK at T2 by Condition (see Supplementary Figure 3), the contrast between False fonts and Syllables changed significantly with varying values of LNK T2 (β = 5.20; SE = 1.13; t = 4.60; p < 0.001), such that the more negative the amplitude for the Syllables, the better the LNK, while the latter did not modulate amplitude values for the False fonts (see Supplementary table 5 in the Appendix for description of the least square means for the different factors' levels).

[Insert Figure 3 around here]

Table 5.

a. Summary of results of the model with amplitude of the EEG at T2 as DV and Condition, Hemisphere, Beat synchronisation T2, LNK T2, Age and IQ as predictors

Model characteristics				
	n	**obs**	**M R^2**	**C R^2**
	17	68	.58	.93
Random effects				
		SD		
By-subject intercept		3.19		
By-subject slope Condition		1.95		
By-subject slope Hemisphere		2.04		
Residual		1.61		
Fixed effects				
	β	**SE**	***t***	***p***

Intercept	-58.98	21.57	-2.735	0.01*
BSynch T2	0.08	0.92	0.08	0.93
Condition[1]	-4.20	0.78	-5.42	<0.001*
Hemisphere[2]	-.49	0.79	-.63	0.54
LNK T2	-.28	2.02	-.14	0.89
Age	0.80	0.29	2.73	0.01*
IQ	1.11	0.38	2.92	0.01*
BSynch T2 * Condition	-1.16	0.65	-1.76	0.09
BSynch T2 * Hemisphere	-.48	0.67	-.71	0.48
Condition * Hemisphere	0.73	0.83	0.88	0.39
LNK T2 * Condition	-4.87	1.34	-3.63	0.001*
LNK T2 * Hemisphere	-.38	1.37	-.28	0.79
BSynch T2 * Condition * Hemisphere	1.49	0.71	2.12	0.05*
LNK T2 * Condition * Hemisphere	-.68	1.44	-.47	0.64

[1] Reference: False fonts; [2] Reference: Left hemisphere. BSynch = Beat synchronisation; C R^2 = Conditional R square; FF = False fonts; M R^2 = Marginal R square; obs = total of observations; Syll = Syllables

b. Tukey HSD post-hoc comparison of the triple interaction Beat synchronisation T2 by Condition by Hemisphere

Contrast	β	**SE**	***t***	***p***
	Left Hemisphere			
False fonts - Syllables	1.15	.66	1.76	.09
	Right Hemisphere			
	-.33	.66	.51	.61

3.3.3. Phonology as a mediator of the rhythm-N170 relation

Regarding T1, the ANOVA comparing the model with and without the Pseudoword repetition scores revealed that the models were not significantly different, suggesting that the phonological short-term memory skills of the children did not mediate in the relation between our behavioural predictors and the amplitude of the N170 in our sample. The upper part of Table 6 shows the detailed output of the ANOVA. Furthermore, the model containing the Pseudoword repetition scores did not capture significantly more variance as compared to the model without them (Model without Pseudoword repetition: marginal R^2= .55; conditional R^2 = .95; Model with Pseudoword repetition: marginal R^2= .47; conditional R^2 = .96).

At T2, the ANOVA showed again that the model including the Pseudoword repetition factor was not significantly different of the model that did not include it (see bottom part of Table 4 for the output of this analysis) and that the marginal variance captured by the new model was even reduced (Model without Pseudoword repetition: marginal R^2= .58; conditional R^2 = .93; Model with Pseudoword repetition: marginal R^2= .44; conditional R^2 = .96).

Table 6. Results of the ANOVAs comparing the models with and without the score in the Pseudoword repetition task at T1 (upper part) and T2 (bottom part)

	Df	**AIC**	**BIC**	**LogL**	**Deviance**	χ^2	***p***
				T1			
Without PWRep T1	17	380.92	418.65	-173.46	346.92		
With PWRep T1	22	384.65	433.48	-170.32	340.65	6.27	0.28
				T2			
Without PWRep T2	21	366.11	411.45	-162.06	324.11		
With PWRep T2	22	367.82	415.32	-161.91	323.82	0.29	0.59

LogL = Log likelihood; PWRep = Pseudoword repetition

4. Discussion

The aim of this study was to assess the relation between pre-readers' rhythm sensitivity and two indexes of reading readiness related to orthographic sensitivity: LNK and the N170 response to print. Our results show that rhythmic abilities are tightly related to both indexes, providing further evidence that rhythmic abilities could be a valid measure to assess reading readiness at very young ages, and therefore to aid in the detection of children at risk of developing reading difficulties.

Rhythm sensitivity and behavioural orthographic sensitivity (Letter name knowledge)

The performance of the children in the Beat synchronisation and in the Letter name knowledge (LNK) tasks was tightly related across the two testing times of the current study, but especially when children were older. At T1, Beat synchronisation and LNK performance did not correlate, while they did notably in T2. This result goes in line with previous proposals suggesting that the development of a unitary construct of reading predictors could be progressive, such that the relation between predictors is expected to remain modest at very young ages (four years old, in our case) and increase with age (e.g. Schatschneider et al., 2004). Moreover, LNK at T1 significantly predicted Beat synchronisation at T2, while Beat synchronisation at T1 predicted LNK at T2, although only marginally ($R = 0.36$; $p = 0.06$). Overall, this evidence confirms our hypothesis that rhythmic abilities are not only related to phonological reading predictors (Lundetræ & Thomson, 2017; Moritz et al., 2013; Carr et al., 2014), but that this relation extends to more orthographic predictors of future reading skills such as letter knowledge. Moreover, our longitudinal data complement previous cross-sectional results showing that rhythm sensitivity predicts letter knowledge in pre-readers (Ozernov-Palchik et al., 2018).

Rhythm sensitivity and neural reading readiness (N170)

Before addressing the question of the relation between rhythm sensitivity and the N170, we will discuss why the negative ERP elicited by the linguistic stimuli in our task is a valid neural measure of print sensitivity in pre-readers. First and most importantly, syllables elicited a negative ERP response as opposed to control stimuli, in line with previous studies (Bach et al., 2013; Brem et al., 2013; Lochy et al., 2016; cf. Brem et al., 2005; Mauer et al., 2006). Second, although the N170 response was different from the typical adult N170 response since it was delayed (around 200 ms in the present study vs. the typical 170 ms latency in adults) and prominently bilateral (Bentin et al., 1999; Brem et al., 2005; Maurer et al., 2006; Maurer, Brandeis, et al., 2005), it is in line with studies that report different topographies and latencies for the same EEG components in children as compared to adults (Grossi, Coch, Coffey-Corina, Holcomb, & Neville, 2001; Holcomb, Coffey, & Neville, 1992; Maurer, Brem, et al., 2005; Taylor & Keenan, 1999). Whereas the absence of a left-hemispheric specialization for print fits with the idea that left-lateralised responses develop after the acquisition of reading (e.g. Schlaggar & McCandliss, 2007; see Shaywitz et al., 2004 for a review), our result contrasts with previous studies in pre-readers that found either a left-lateralised (e.g. Bach et al., 2013) or a right-lateralised response to print in pre-readers (Maurer, Brem, et al., 2005). We cannot discard that these inconsistencies might be due to the use of distinct paradigms across studies. Overall, we believe that our novel passive paradigm was able to capture N170 responses reflecting pre-reading neural print sensitivity (or *linguistic awareness* for print). In addition, it is noteworthy that this paradigm was particularly suited for young children since it required no memory load, low visual perceptual demands (consonant-vowel syllables), and little time to be completed (approximately four minutes).

One of the main questions of the present study was whether this early neural marker of orthographic sensitivity could be predicted by rhythm sensitivity even before children started to be taught formally how to read. We showed that Beat synchronisation performance measured at four years old was able to predict N170 responses one year later, such that children with better rhythmic abilities at T1 showed a larger negativity only for syllables and in the left hemisphere at T2. This relation between rhythmic sensitivity and N170 responses was not only found across phases, but also within T2, although this result did not survive the *post-hoc* correction for multiple comparisons. Overall, our results show that performance on the Beat synchronisation task might be a good index of future neural sensitivity to orthographic precursors of reading in pre-readers.

It was interesting to see that LNK knowledge did not predicted N170 responses one year later (whereas Beat synchronization did). This is at odds with our hypothesis that both LNK and N170 reflect two measures of one same construct, namely pre-reading orthographic sensitivity. Nevertheless, at T2, LNK explained variance in N170 response to syllables in both hemispheres, beyond the variance explained by Beat synchronisation. These results are in line with studies showing that stable and consistent LNK performance is reported no earlier than five years of age (Mason, 1980; McBride-Chang, 1999; Treiman et al., 2008; Worden & Boettcher, 1990). This would support our result that variance of N170 responses were best captured by LNK later in development (T2). Interestingly, Beat synchronization predicted N170 responses from T1 on, possibly because this ability develops very early in life and is relatively well established at three or four years of age (Drake et al., 2000; Droit-Volet, 1998; Fraisse, Pichot, & Clairouin, 1969; McAuley, Jones, Holub, Johnston, & Miller, 2006; Pouthas, Droit, Jacquet, & Wearden, 1990; Provasi & Bobin-Bègue, 2003).

The mediation of phonology-related skills in the rhythm-orthographic sensitivity relation

Our results showed that the relation between rhythmic abilities and our behavioural and neural measures of pre-reading orthographic sensitivity was not mediated by the phonological short-term memory skills of the children. In principle, this result is at odds with recent theories of developmental reading disorders claiming that the rhythm sensitivity-reading relation should be mediated by phonological abilities (Goswami, 2011; Lallier et al., 2017; Moritz et al., 2013). Nevertheless, it is in line with studies showing that the phonological skills of the children might not be able to explain the relation between rhythm sensitivity and letter knowledge in pre-readers (Ozernov-Palchik et al., 2018), and also with a recent research line suggesting that the contribution of rhythmic abilities to reading at very early stages might not be mediated by phonological skills (e.g. Calet et al., 2015; Calet et al., 2019; González-Trujillo et al., 2014; Holliman et al., 2010). Some have suggested that this non-phonological route could be rooted in the development of written prosody knowledge such as punctuation marks (Calet et al., 2019). In fact, it is noteworthy that the orthographic lexical accent (i.e., “tilde”) is an important orthographic prosodic feature of Spanish, one of the languages known by the children of the present study. Interestingly, the development of pre-reading written prosody knowledge might rely on the implicit extraction of prosodic visual rules from “out-of-school” home literacy experiences, whose effect on reading development was shown to be relatively independent from phonological skills (see below). However, we can barely speculate on this point at present since our study was not designed to directly address this issue and no study to date has examined this question.

In addition, it is important to note that the only phonological measure available for our analyses was phonological short-term memory, whereas previous studies have mostly used phoneme awareness. In fact, we tried, but could not analyse the phoneme awareness data due to floor effects. Interestingly, the fact that children were not able to complete the phonemic awareness task in pre-reading stages is in line with the idea that the relation between phonological skills and reading acquisition is reciprocal rather than unidirectional (Castles & Coltheart, 2004; Stuart & Coltheart, 1988), and goes in line with proposals suggesting that the development of phonological skills and orthographic sensitivity might be independent at very early reading stages (Wood, 2004). A follow-up analysis testing the relation between letter knowledge and phonological short-term memory in our sample showed that these were not related at any of the testing times

(T1: *n* = 41, *r* = .18; T2: *n* = 34, *r* = .25) suggesting that, at pre-reading stages, these skills might be still rather independent (Wood, 2004). As noted before, children may develop rudimentary phonological skills from birth through the perception of speech, whereas early orthographic and print sensitivity would develop through a possibly independent pathway rooted in inter-individual differences regarding the exposure to written material in pre-reading stages. Accordingly, Evans, Shaw & Bell (2000) showed that home activities involving letters with pre-readers predicted their LNK but not their phonological sensitivity. In addition, Levy, Gong, Hessels, Evans, & Jared (2006) showed that home literacy experiences did not predict the phonological sensitivity of 4 years-old children. Overall, this could support the proposal that during pre-reading stages, the contribution of rhythm sensitivity to orthographic sensitivity and to phonological sensitivity may follow independent developmental pathways.

Limitations of the present study and future directions

We are aware that the rationale supporting the hypothesis that rhythm sensitivity could predict orthographic precursors of reading independently of phonological skills is still weak and needs further theoretical investigation. Since our study was the first testing the relation between rhythm and both behavioural and brain orthographic sensitivity in very young children, further research is needed to conclude on the possible mediation of phonological and non-phonological factors in this relation. Also, we acknowledge that only one phonological index was available in the present study, limiting our conclusions on these aspects. Future studies should thrive collecting a large range of phonological (i.e., phonological awareness, phonological short-term memory, phonological access) as well as potential non-phonological (e.g., written prosody or home literacy experiences) measures to characterize further the relationship between rhythm sensitivity and reading development. Lastly, it is important to note that our sample size was reduced, especially for the analyses including EEG data. Although we tried to compensate for this using a robust statistical method such as mixed effects models, further studies with larger sample sizes are needed to draw more definite conclusions on the results reported here.

Conclusions

Our results show that pre-readers’ rhythmic sensitivity as assessed in a beat synchronization task could be an early marker of behavioural (LNK) and neurological (N170 amplitude) orthographic sensitivity in children that have still not received formal reading instruction. Importantly, these relations were not mediated by the phonological short-term memory skills of the children. Thanks to its longitudinal design, our study offers new insights into the power that rhythm sensitivity might have to predict future brain activity related to reading, specifically the neural sensitivity to print. Due to the difficulties of testing young children with neuropsychological techniques, discovering behavioural markers that predict neural activity at later stages of reading development is of high scientific and clinical interest. Further research confirming the robustness of this relation could offer a child-friendly way to predict sensitivity to print in children, and hence increase the probability of detecting children at risk of future reading problems.

Funding

This work was supported by the Severo Ochoa Program grant [SEV-2015-049]; ML was supported by the Spanish Ministry of Economy and Competitiveness [RTI2018-096242-B-I00, Ramón y Cajal Fellowship]; NL was supported by the Spanish Ministry of Economy and Competitiveness [PSI2015-65694-P], the Basque Government [PI_2016_1_0014], the European Commission 7th Framework Programme (AThEME project) and the European Research Council (ERC-2011-ADG-295362).

Conflict of interest statement

Declarations of interest: none.

Acknowledgements

The authors wish to thank Alexia Antzaka for her help with testing the children and her insightful comments on data analysis, Sara Aurtenetxe for providing the false font materials and the children and their parents for participating in this study.

References

1. Baayen, R. H., Davidson, D. J., & Bates, D. M. (2008). Mixed-effects modeling with crossed random effects for subjects and items. *Journal of Memory and Language*, *59*(4), 390–412. https://doi.org/10.1016/j.jml.2007.12.005

2. Bach, S., Richardson, U., Brandeis, D., Martin, E., & Brem, S. (2013). Print-specific multimodal brain activation in kindergarten improves prediction of reading skills in second grade. *NeuroImage*, *82*, 605–615. http://doi.org/10.1016/j.neuroimage.2013.05.062

3. Barr, D. J., Levy, R., Scheepers, C., & Tily, H. J. (2013). Random effects structure for confirmatory hypothesis testing: Keep it maximal. *Journal of Memory and Language*, *68*(3), 255–278. https://doi.org/10.1016/j.jml.2012.11.001

4. Bell, A. J., & Sejnowski, T. J. (1995). An information-maximization approach to blind separation and blind deconvolution. *Neural computation*, *7*(6), 1129-1159. https://doi.org/10.1162/neco.1995.7.6.1129

5. Bentin, S., Mouchetant-Rostaing, Y., Giard, M. H., Echallier, J. F., & Pernier, J. (1999). ERP manifestations of processing printed words at different psycholinguistic levels: Time course and scalp distribution. *Journal of Cognitive Neuroscience*, *11*(3), 235–260. http://doi.org/10.1162/089892999563373

6. Berens, P. (2009). CircStat: a MATLAB toolbox for circular statistics. *Journal of Statistical Software*, *31*(10), 1-21. https://doi.org/10.18637/jss.v031.i10

7. Bishop, D. V. M., & Snowling, M. J. (2004). Developmental dyslexia and specific language impairment: Same or different?. *Psychological Bulletin, 130*(6), 858–886. https://doi.org/10.1037/0033-2909.130.6.858

8. Brem, S., Bach, S., Kujala, J. V., Maurer, U., Lyytinen, H., Richardson, U., & Brandeis, D. (2013). An electrophysiological study of print processing in kindergarten: The contribution of the visual N1 as a predictor of reading outcome. *Developmental Neuropsychology*, *38*(8), 567–594. http://doi.org/10.1080/87565641.2013.828729

9. Brem, S., Halder, P., Bucher, K., Summers, P., Martin, E., & Brandeis, D. (2009). Tuning of the visual word processing system: Distinct developmental ERP and fMRI effects. *Human Brain Mapping*, *30*(6), 1833–1844. http://doi.org/10.1002/hbm.20751

10. Brem, S., Lang-Dullenkopf, A., Maurer, U., Halder, P., Bucher, K., & Brandeis, D. (2005). Neurophysiological signs of rapidly emerging visual expertise for symbol strings. *Neuroreport*, *16*(1), 45–8. http://doi.org/00001756-200501190-00011

11. Byrne, B., & Fielding-Barnsley, R. (1990). Acquiring the alphabetic principle: A case for teaching recognition of phoneme identity. *Journal of Educational Psychology*, *82*(4), 805–812. https://doi.org/10.1037/0022-0663.82.4.805

12. Calet, N., Gutiérrez-Palma, N., Simpson, I. C., González-Trujillo, M. C., & Defior, S. (2015). Suprasegmental Phonology Development and Reading Acquisition: A Longitudinal Study. *Scientific Studies of Reading*, *19*(1), 51–71. https://doi.org/10.1080/10888438.2014.976342

13. Carr, K. W., 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 Science, 111*(40), 14559-14564. https://doi.org/10.1073/pnas.1406219111

14. Castles, A., & Coltheart, M. (2004). Is there a causal link from phonological awareness to success in learning to read? *Cognition*, *91*(January 2003), 77–111. https://doi.org/10.1016/S0010-0277(03)00164-1

15. Corriveau, K. H., & Goswami, U. (2009). Rhythmic motor entrainment in children with speech and language impairments: Tapping to the beat. *Cortex, 45*(1), 119–130. https://doi.org/10.1016/j.cortex.2007.09.008

16. Corriveau, K. H., Goswami, U., & Thomson, J. M. (2010). Auditory processing and early literacy skills in a preschool and kindergarten population. *Journal of Learning Disabilities*, *43*, 369–382. https://doi.org/10.1177/0022219410369071

17. Drake, C., Jones, M. R., & Baruch, C. (2000). The development of rhythmic attending in auditory sequences: Attunement, referent period, focal attending. *Cognition*, *77*(3), 251-288. https://doi.org/10.1016/S0010-0277(00)00106-2

18. Droit-Volet, S. (1998). Time estimation in young children: An initial force rule governing time production. *Journal of Experimental Child Psychology*, *68*(3), 236–249. http://doi.org/10.1006/jecp.1997.2430

19. Eberhard-Moscicka, A. K., Jost, L. B., Raith, M., & Maurer, U. (2015). Neurocognitive mechanisms of learning to read: Print tuning in beginning readers related to word-reading fluency and semantics but not phonology. *Developmental Science*, *18*(1), 106–118. http://doi.org/10.1111/desc.12189

20. Evans, M. A., Bell, M., Shaw, D., Moretti, S., & Page, J. (2006). Letter names, letter sounds and phonological awareness: An examination of kindergarten children across letters and of letters across children. *Reading and Writing, 19*(9), 959-989. https://doi.org/10.1007/s11145-006-9026-x

21. Evans, M. A., Shaw, D., & Bell, M. (2000). Home literacy activities and their influence on early literacy skills. *Canadian Journal of Experimental Psychology/Revue canadienne de psychologie expérimentale, 54*(2), 65.

22. Fisher, N. I. (1993). *Statistical analysis of circular data*. Cambridge: Cambridge University Press. https://doi.org/10.1017/CBO9780511564345

23. Flaugnacco, E., Lopez, L., Terribili, C., Zoia, S., Buda, S., Tilli, S., … Schön, D. (2014). Rhythm perception and production predict reading abilities in developmental dyslexia. *Frontiers in Human Neuroscience, 8,* 392. https://doi.org/10.3389/fnhum.2014.00392

24. Fraisse, P., Pichot, P., & Claiouin, G. (1969). Les aptitudes rhythmiques. Etude comparée des oligoprhènes et des enfants normaux. *Journal de Psychologie Normale et Pathologique, 42,* 309-330

25. González-Trujillo, M. C., Defior, S., & Gutiérrez-Palma, N. (2014). The role of nonspeech rhythm in Spanish word reading. *Journal of Research in Reading, 37*(3), 316–330. https://doi.org/10.1111/j.1467-9817.2012.01529.x

26. Goswami, U. (2011). A temporal sampling framework for developmental dyslexia. *Trends in Cognitive Sciences*, *15*(1), 3–10. https://doi.org/10.1016/j.tics.2010.10.001

27. Goswami, U., Power, A. J., Lallier, M., & Facoetti, A. (2014). Oscillatory "temporal sampling" and developmental dyslexia: toward an over-arching theoretical framework. *Frontiers in Human Neuroscience, 8*, 904. https://doi.org/10.3389/fnhum.2014.00904

28. Grossi, G., Coch, D., Coffey-Corina, S., Holcomb, P. J., & Neville, H. J. (2001). Phonological processing in visual rhyming: a developmental ERP study. *Journal of Cognitive Neuroscience, 13*(5), 610–625. http://doi.org/10.1162/089892901750363190

29. Holcomb, P. J., Coffey, S., & Neville, H. (1992). Visual and auditory sentence processing: A developmental analysis using event-related brain potentials. *Developmental Neuropsychology, 8*(2-3), 203–241. http://doi.org/10.1080/87565649209540525

30. Holliman, A., Critten, S., Lawrence, T., Harrison, E., Wood, C., & Hughes, D. (2014). Modeling the relationship between prosodic sensitivity and early literacy. *Reading Research Quarterly*, *49*(4), 469–482. https://doi.org/10.1002/rrq.82

31. Holliman, A. J., Wood, C., & Sheehy, K. (2008). Sensitivity to speech rhythm explains individual differences in reading ability independently of phonological awareness. *British Journal of Developmental Psychology*, *26*(3), 357–367. https://doi.org/10.1348/026151007X241623

32. Holliman, A. J., Wood, C., & Sheehy, K. (2010). Does speech rhythm sensitivity predict children's reading ability 1 year later? *Journal of Educational Psychology*, *102*(2), 356–366. https://doi.org/10.1037/a0018049

33. Huss, M., Verney, J. P., Fosker, T., Mead, N., & Goswami, U. (2011). Music, rhythm, rise time perception and developmental dyslexia: Perception of musical meter predicts reading and phonology. *Cortex*, *47*(6), 674–689. https://doi.org/10.1016/j.cortex.2010.07.010

34. Jorm, A. F. (1979). The cognitive and neurological basis of developmental dyslexia: A theoretical framework and review. *Cognition*, *7*(1), 19–33. http://doi.org/10.1016/0010-0277(79)90008-8

35. Kirschner, S., & Tomasello, M. (2009). Joint drumming: Social context facilitates synchronization in preschool children. *Journal of Experimental Child Psychology*, *102*(3), 299–314. https://doi.org/10.1016/j.jecp.2008.07.005

36. Lallier, M., Lizarazu, M., Molinaro, N., Bourguignon, M., Ríos-López, P., & Carreiras, M. (2018). From auditory rhythm processing to grapheme-to-phoneme conversion: How neural oscillations can shed light on developmental dyslexia. In *Reading and Dyslexia* (pp. 141-157). Cham: Springer. https://doi.org/10.1007/978-3-319-90805-2_8

37. Lallier, M., Molinaro, N., Lizarazu, M., Bourguignon, M., & Carreiras, M. (2017). Amodal atypical neural oscillatory activity in dyslexia. *Clinical Psychological Science*, *5*(2), 379–401. https://doi.org/10.1177/2167702616670119

38. Lenth, R. V. (2016). Least-Squares Means: The *R* Package lsmeans. *Journal of Statistical Software*, *69*(1), 1-33. https://doi.org/10.18637/jss.v069.i01

39. Leong, V., Hämäläinen, J., Soltész, F., & Goswami, U. (2011). Rise time perception and detection of syllable stress in adults with developmental dyslexia.

Journal of Memory and Language, *64*(1), 59–73. https://doi.org/10.1016/j.jml.2010.09.003

40. Levy, B. A., Gong, Z., Hessels, S., Evans, M. A., & Jared, D. (2006). Understanding print: Early reading development and the contributions of home literacy experiences. *Journal of Experimental Child Psychology, 93*(1), 63-93.

41. Lochy, A., Van Reybroeck, M., & Rossion, B. (2016). Left cortical specialization for visual letter strings predicts rudimentary knowledge of letter-sound association in preschoolers. *Proceedings of the National Academy of Sciences, 113*(30), 8544–8549. http://doi.org/10.1073/pnas.1520366113

42. Lundetræ, K., & Thomson, J. M. (2017). Rhythm production at school entry as a predictor of poor reading and spelling at the end of first grade. *Reading and Writing*, *31*(1), 215-237. https://doi.org/10.1007/s11145-017-9782-9

43. Mason, J. M. (1980). When do children begin to read: An exploration of four year old children's letter and word reading competencies. *Reading Research Quarterly*, *15*(2), 203-227. http://doi.org/10.2307/747325

44. Maurer, U., Brandeis, D., & McCandliss, B. D. (2005). Fast, visual specialization for reading in English revealed by the topography of the N170 ERP response. *Behavioral and Brain Functions*, *1*(13), 1–12. http://doi.org/10.1186/1744-9081-1-13

45. Maurer, U., Brem, S., Bucher, K., & Brandeis, D. (2005). Emerging neurophysiological specialization for letter strings. *Journal of Cognitive Neuroscience*, *17*(10), 1532–1552. http://doi.org/10.1162/089892905774597218

46. Maurer, U., Brem, S., Kranz, F., Bucher, K., Benz, R., Halder, P., … Brandeis, D. (2006). Coarse neural tuning for print peaks when children learn to read. *NeuroImage*, *33*(2), 749–758. http://doi.org/10.1016/j.neuroimage.2006.06.025

47. McAuley, J. D., Jones, M. R., Holub, S., Johnston, H. M., & Miller, N. S. (2006). The time of our lives: Life span development of timing and event tracking. *Journal of Experimental Psychology: General*, *135*(3), 348–367. https://doi.org/10.1037/0096-3445.135.3.348

48. McBride-Chang, C. (1999). The ABCs of the ABCs: The development of letter-name and letter-sound knowledge. *Merrill-Palmer Quarterly, 45*(2), 285-308. Retrieved from http://www.jstor.org/stable/23093679

49. Moritz, C., Yampolsky, S., Papadelis, G., Thomson, J., & Wolf, M. (2013). Links between early rhythm skills, musical training, and phonological awareness. *Reading and Writing, 26,* 739–769. https://doi.org/10.1007/s11145-012-9389-0

50. Muter, V., & Snowling, M. (1998). Concurrent and longitudinal predictors of reading: The role of metalinguistic and short-term memory skills. *Reading Research Quarterly, 33*(3), 320-337. https://doi.org/10.1598/RRQ.33.3.4

51. Nozaradan, S., Zerouali, Y., Peretz, I., & Mouraux, A. (2015). Capturing with EEG the Neural Entrainment and Coupling Underlying Sensorimotor Synchronization to the Beat. *Cerebral Cortex*, *25*(3), 736–747. https://doi.org/10.1093/cercor/bht261

52. 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. http://doi.org/10.1016/j.jecp.2017.11.009

53. Peirce, J. W. (2009). Generating stimuli for neuroscience using PsychoPy. *Frontiers in Neuroinformatics*, *2*, 10. https://doi.org/10.3389/neuro.11.010.2008

54. Pouthas, V., Droit, S., Jacquet, a Y., & Wearden, J. H. (1990). Temporal differentiation of response duration in children of different ages: developmental changes in relations between verbal and nonverbal behavior. *Journal of the Experimental Analysis of Behavior*, *53*(1), 21–31. http://doi.org/10.1901/jeab.1990.53-21

55. Provasi, J., & Bobin-Bègue, A. (2003). Spontaneous motor tempo and rhythmical synchronisation in 21/2- and 4-year-old children. *International Journal of Behavioral Development*, *27*(3), 220–231. http://doi.org/10.1080/01650250244000290

56. R Core Team (2017). R: A language and environment for statistical computing. R Foundation for Statistical Computing. Vienna, Austria. URL https://www.R-project.org/

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

58. Richardson, U., Thomson, J. M., Scott, S. K., & Goswami, U. (2004). Auditory processing skills and phonological representation in dyslexic children. *Dyslexia, 10*(3), 215–233. https://doi.org/10.1002/dys.276

59. Rogerson, P. A. (2001). A statistical method for the detection of geographic clustering. *Geographical Analysis*, *33*(3), 215–227. http://doi.org/10.1111/j.1538-4632.2001.tb00445.x

60. Schatschneider, C., Fletcher, J. M., Francis, D. J., Carlson, C. D., & Foorman, B. R. (2004). Kindergarten prediction of reading skills: A longitudinal comparative analysis. *Journal of Educational Psychology, 96*(2), 265–282. https://doi.org/10.1037/0022-0663.96.2.265

61. Schlaggar, B. L., & McCandliss, B. D. (2007). Development of Neural Systems for Reading. *Annual Review of Neuroscience, 30*(1), 475–503. http://doi.org/10.1146/annurev.neuro.28.061604.135645

62. Shaywitz, B. A., Shaywitz, S. E., Blachman, B., Pugh, K. R., Fulbright, R. K., Skudlarski, P., … Gore, J. C. (2004). Development of left occipitotemporal systems for skilled reading in children after a phonologically-based intervention. *Biological Psychiatry*, *55*(7), 685–691. http://doi.org/10.1016/j.biopsych.2004.01.006

63. Snowling, M. J. (2001). From language to reading and dyslexia. *Dyslexia, 7*(1), 37–46. https://doi.org/10.1002/dys.185

64. Stanovich, K. E., Cunningham, A. E., & Cramer, B. B. (1984). Assessing phonological awareness in kindergarten children: Issues of task comparability. *Journal of Experimental Child Psychology, 38*(2), 175–190. https://doi.org/10.1016/0022-0965(84)90120-6

65. Stuart, M., & Coltheart, M. (1988). Does reading develop in a sequence of stages? *Cognition*, *30*(2), 139–181. https://doi.org/10.1016/0010-0277(88)90038-8

66. Taylor, M. J., & Keenan, N. K. (1999). ERPs to orthographic, phonological, and semantic tasks in dyslexic children with auditory processing impairment. *Developmental Neuropsychology*, *15*(2), 307–326. http://doi.org/10.1080/87565649909540751

67. Thomson, J. M., Fryer, B., Maltby, J., & Goswami, U. (2006). Auditory and motor rhythm awareness in adults with dyslexia. *Journal of Research in Reading*, *29*(3), 334–348. https://doi.org/10.1111/j.1467-9817.2006.00312.x

68. Thomson, J. M., & Goswami, U. (2008). Rhythmic processing in children with developmental dyslexia: Auditory and motor rhythms link to reading and spelling. *Journal of Physiology-Paris*, *102*(1-3), 120–129. https://doi.org/10.1016/j.jphysparis.2008.03.007

69. Tierney, A. T., & Kraus, N. (2013). The ability to tap to a beat relates to cognitive, linguistic, and perceptual skills. *Brain and Language*, *124*, 225–231. https://doi.org/10.1016/j.bandl.2012.12.014

70. Treiman, R., Tincoff, R., Rodriguez, K., Mouzaki, A., & Francis, D. J. (2008). The foundations of literacy: Learning the sounds of letters. *Child Development*, *69*(6), 1524-1540. https://doi.org/10.1111/j.1467-8624.1998.tb06175.x

71. Wechsler, D. (2002). *The Wechsler Preschool and Primary Scale of Intelligence* (3rd ed.) *(WPPSI-III)*. San Antonio: The Psychological Corporation.

72. Wood, C. (2004). Do levels of pre-school alphabetic tuition affect the development of phonological awareness and early literacy? *Educational Psychology*, *24*(1), 3–11. https://doi.org/10.1080/0144341032000146403

73. Wood, C., Wade-Woolley, L., & Holliman, A. J. (2009). Phonological awareness: Beyond phonemes. In C. Wood & V. Connelly (Eds.), Contemporary perspectives on reading and spelling (pp. 7–23). New York: Routledge.

74. Wood, C., & Terrell, C. (1998). Poor readers' ability to detect speech rhythm and perceive rapid speech. *British Journal of Developmental Psychology*, *16*(3), 397–413. https://doi.org/10.1111/j.2044-835X.1998.tb00760.x

75. Worden, P. E., & Boettcher, W. (1990). Young children's acquisition of alphabet knowledge. *Journal of Literacy Research*, *22*(3), 277–295. http://doi.org/10.1080/10862969009547711

76. Zuur, A. F., Ieno, E. N., & Elphick, C. S. (2010). A protocol for data exploration to avoid common statistical problems. *Methods in Ecology and Evolution*, *1*(1), 3–14. https://doi.org/10.1111/j.2041-210X.2009.00001.x

Appendix

1. Supplementary analysis: Syllable versus False font analysis in the individual electrodes of interest

The FDR-corrected multiple point-by-point paired *t*-tests analysis in the window between -200 ms and 190 ms confirmed that there was no difference between the waves neither in any time point nor in any of our electrodes of interest (all *p*s > 0.05). For our window of interest (from 190 to 325), we found that the waves started being significantly different at variable time points depending on electrode and hemisphere. In left posterior electrodes the significant effect between conditions started in a range between 194 and 204 ms (P3onset: 204 ms; P7 onset: 194 ms; O1 onset: 200 ms; all *p*s < 0.05), while in right electrodes it started later, in a range between 218 and 236 ms (P4 onset: 236 ms; P8 onset: 218 ms; O2 onset: 222 ms; all *p*s < 0.05). Regardless of electrode, the difference between waves was sustained until the end of our window of interest (325 ms; all *p*s < 0.05). Although the effect between conditions was apparently still present after 325 ms, we deemed that this could reflect a carry-over of the differences observed in the N170 interval and did not consider for further analysis any activity occurring after this time point (see main text for full justification).

2. Supplementary results

Due to the longitudinal character of our design and to the high rate of experimental mortality in studies testing young children, the data used to feed the linear mixed models including only the EEG was different to the data used to feed the models including both the EEG and the behavioural tasks. Hence, our models were not directly comparable in statistical terms. This led us to fit two supplementary models for the EEG data including only the children that had also completed the behavioural tasks in T1 and in T2 (n = 17) to obtain statistical support for the fact that adding our behavioural data increased the model's captured variance by the original EEG models for T1 and T2, respectively. Regarding the selection of the fixed effects of the optimal model, we used maximum likelihood (ML) estimation to fit the models to be able to compare across fits with identical random effects but different nested fixed effects (Zuur, Ieno, & Elphick, 2010). Model selection was then based on visual inspection of the coefficients and of the Akaike information criterion (AIC), and on the results of a χ^2 test across models.

Supplementary Table 1 shows the output of the new EEG-only model for T1 (n = 17, total observations = 68) for both the fixed and the random effect. Despite the coefficients change due to the difference in the sample size and the number of observations, the effects significance was identical to the original EEG model, so they will not be discussed further. Our interest with this analysis was to obtain statistical support for the hypothesis that adding our behavioural predictors at T1 improved the model's fit, and hence we proceeded to compute an analysis of variance (χ^2 test) to compare the models without and with the behavioural predictors at T1. The result of this analysis confirmed that the models differed significantly (new EEG model: AIC = 359.94, deviance = 333.94; EEG + behavioural model: AIC = 357.25, deviance = 315.25; $\chi^2 = 18.69$, $p = 0.02$). As indicated by the AIC values and by the visual inspection of the coefficients of the two models, the model including the behavioural data seemed to represent a better fit of the data as compared to the new EEG model.

Supplementary table 1. Output of the EEG model including only the children that also completed the behavioural tasks at T1

Fixed effects structure	Estimate (β)	SE	*t*	*P*
Intercept	-45.88	17.68	-2.60	0.02
Condition (FF [reference] vs Syllables)	-5.80	1.03	-5.63	<0.001*
Hemisphere (Left [reference] vs. Right)	-1.20	0.80	-1.50	0.14
Age	0.60	0.24	2.47	0.03
IQ	1.30	0.40	3.28	<0.01
Condition * Hemisphere	1.51	0.83	1.82	0.09

Random effects' *SD*	By-subject Intercept	3.32
	By-subject slope for	3.38

	Condition			
	By-subject slope for Hemisphere	2.17		
	Residual	1.66		

FF = False fonts

We followed the same procedure for the model comparison at T2. The output of the new EEG model for T2 (n = 17, total observations = 68) can be seen in Supplementary table 2. Again, the effects significance was identical to the original EEG model, so they will be not further discussed. The χ^2 test comparing the models without and with the behavioural predictors at T2 confirmed that these models also differed significantly (new EEG model: AIC = 385.93, deviance = 359.93; behavioural + EEG model: AIC = 376.37, deviance = 334.37; $\chi^2 = 25.55$, $p = 0.001$), such that the model with the behavioural predictors improved the data's fit.

Supplementary table 2. Output of the EEG model including only the children that also completed the behavioural tasks at T2

Fixed effects structure	Estimate (β)	SE	*t*	*P*
Intercept	-60.07	18.08	-3.32	<0.01
Condition (FF [reference] vs Syllables)	-5.53	1.10	-5.02	<0.001*
Hemisphere (Left [reference] vs. Right)	-0.69	0.78	-0.89	0.38
Age	0.81	0.25	3.23	<0.01
IQ	1.12	0.33	3.42	<0.01
Condition * Hemisphere	0.96	0.88	1.09	0.29

Random effects' *SD*	By-subject Intercept	3.13		
	By-subject slope for Condition	3.75		
	By-subject	1.91		

	slope for Hemisphere	
	Residual	1.82

FF = False fonts

Supplementary table 3. Least square means for the interaction Beat synchronization T1 by Condition by Hemisphere based on pairwise contrasts between the levels of Condition (Syllables and False fonts) and Hemisphere (Left and Right) for varying trends of Beat synchronization T1. Group numbers indicate the group of the slope. Different slope groups indicate significant differences between the slopes after multiplicity adjustment (Tukey HSD).

Left hemisphere

Condition	BSynch T1 trend	SE	Lower CL	Upper CL	Group
Syllables	-4.24	1.66	-7.72	-0.76	1*
False fonts	-0.73	1.45	-3.76	2.31	2

Right hemisphere

Condition	BSynch T1 trend	SE	Lower CL	Upper CL	Group
Syllables	-1.88	2.14	-6.38	2.61	1
False fonts	-125°	2.12	-5.70	3.20	1

* Significant difference at alpha = 0.05

BSynch = Beat synchronization

Supplementary table 4. Least square means for the interaction Beat synchronisation T2 by Condition by Hemisphere based on pairwise contrasts between the levels of Condition (Syllables and False fonts) and Hemisphere (Left and Right) for varying trends of Beat synchronisation T2. Group numbers indicate the group of the slope. Different slope groups indicate significant differences between the slopes after multiplicity adjustment (Tukey HSD).

Left hemisphere

Condition	BSynch T2 trend	SE	Lower CL	Upper CL	Group
Syllables	-1.08	0.96	-3.10	0.93	1
False fonts	-0.08	0.93	-1.87	2.02	1

Right hemisphere

Condition	BSynch T2 trend	SE	Lower CL	Upper CL	Group
Syllables	-0.40	1.14	-2.78	1.98	1
False fonts	-0.07	1.05	-2.26	2.13	1

BSynch = Beat synchronization

Supplementary table 5. Least square means for the interaction Letter name knowledge T2 by Condition based on pairwise contrasts between the levels of Condition (Syllables and False fonts) for varying trends of Letter name knowledge T2. Group numbers indicate the group of the slope. Different slope groups indicate significant differences between the slopes after multiplicity adjustment (Tukey HSD)

Condition	LNK T2 trend	SE	Lower CL	Upper CL	Group
Syllables	-5.67	2.07	-10.00	-1.34	1*
False fonts	-0.46	2.13	-4.93	4.00	2

LNK = Letter name knowledge