Check for updates

Non-game like training benefits spoken foreign-language processing in children with dyslexia

Katja Junttila[1]*, Anna-Riikka Smolander[1,2], Reima Karhila[3], Mikko Kurimo[3] and Sari Ylinen[1,2]

[1]Cognitive Brain Research Unit, Department of Psychology and Logopedics, Faculty of Medicine, University of Helsinki, Helsinki, Finland, [2]Logopedics, Welfare Sciences, Faculty of Social Sciences, Tampere University, Tampere, Finland, [3]Department of Signal Processing and Acoustics, Aalto University, Espoo, Finland

Children with dyslexia often face difficulties in learning foreign languages, which is reflected as weaker neural activation. However, digital language-learning applications could support learning-induced plastic changes in the brain. Here we aimed to investigate whether plastic changes occur in children with dyslexia more readily after targeted training with a digital language-learning game or similar training without game-like elements. We used auditory event-related potentials (ERPs), specifically, the mismatch negativity (MMN), to study learning-induced changes in the brain responses. Participants were 24 school-aged Finnish-speaking children with dyslexia and 24 age-matched typically reading control children. They trained English speech sounds and words with "Say it again, kid!" (SIAK) language-learning game for 5 weeks between ERP measurements. During the game, the players explored game boards and produced English words aloud to score stars as feedback from an automatic speech recognizer. To compare the effectiveness of the training type (game vs. non-game), we embedded in the game some non-game levels stripped of all game-like elements. In the dyslexia group, the non-game training increased the MMN amplitude more than the game training, whereas in the control group the game training increased the MMN response more than the non-game training. In the dyslexia group, the MMN increase with the non-game training correlated with phonological awareness: the children with poorer phonological awareness showed a larger increase in the MMN response. Improved neural processing of foreign speech sounds as indicated by the MMN increase suggests that targeted training with a simple application could alleviate some spoken foreign-language learning difficulties that are related to phonological processing in children with dyslexia.

1. Introduction

Developmental dyslexia is manifested as reading and writing impairment despite sufficient opportunity for instruction (Bishop and Snowling, 2004). The prevalence of dyslexia ranges from 5 to 20%, depending on definition (Wagner et al., 2020). In addition to reading and writing difficulties, dyslexia is often associated with difficulties in verbal short-term memory (Szenkovits and Ramus, 2005), rapid automatized naming (Bexkens et al., 2015), and phonological processing (Ramus et al., 2013). The difficulties in phonological processing may also be linked to impaired auditory processing of speech (Gu and Bi, 2020), yet the findings have been controversial in this respect. Nevertheless, children with dyslexia typically have difficulties learning spoken novel words (Adlof et al., 2021) and they struggle especially with phonologically demanding tasks which require detecting subtle differences between spoken words (Alt et al., 2017). Consequently, dyslexia does not only impair reading and writing in the native language, but it often hampers learning foreign languages (Downey et al., 2000; Helland and Kaasa, 2005; Soroli et al., 2010; Ylinen et al., 2019).

In our previous study (Ylinen et al., 2019), we found a link between native-language literacy skills and foreign-language word representations in the brain, suggesting that typical readers may benefit from classroom teaching and exposure more than dyslexic readers. A recent study found a hybrid technology game useful for children with dyslexia for foreign language receptive vocabulary learning (Eden and Shmila, 2022). The hybrid technology used by Eden and Shmila (2022) focused on spelling and translation. However, it remains unexplored whether more targeted training techniques focusing on listening to and producing speech could support spoken foreign-language learning in dyslexia (see Jarsve and Tsagari, 2022) and, in addition, which features of targeted training are particularly beneficial for them. Since children with dyslexia have reading difficulties, speech-based training can be expected to be more suitable for their foreign-language learning than reading-based training. Nevertheless, it is not clear which other features of targeted foreign-language training are particularly beneficial for children with dyslexia because studies on foreign-language learning interventions are scarce. For example, our earlier study shows that game-based learning supports the learning of foreign speech sounds more than the use of a non-game application in typically developing children (Junttila et al., 2022), raising the question whether this holds also for children with dyslexia. Digital language-learning games can enhance player's motivation to learn by being easy to use and providing reward and feedback, right amount of challenge, control and autonomy, clearly defined goals, and interactivity (Acquah and Katz, 2020). Therefore, they could be an attractive way for children with dyslexia to learn foreign languages. However, it is not known whether digital language-learning games induce plastic changes in the brain and support the establishment of long-term memory representations similarly for children with dyslexia as for typically reading children. Therefore, this study aimed to determine whether targeted game-based training supports foreign-language learning in dyslexia. Specifically, we compared the effectiveness of playing a digital language-learning game and rehearsing in a non-game application in children with dyslexia.

The game we used focuses on spoken foreign-language learning and it is based on listening and producing speech (Karhila et al., 2017; Ylinen and Kurimo, 2017; Ylinen et al., 2021; Junttila et al., 2022). In the game, players hear English words and produce them aloud, which activates the connections between temporal and frontal language areas that have been found to be deficient in people with dyslexia (Boets et al., 2013; Wang et al., 2017; see Vandermosten et al., 2012 for a review). Speech training with auditory-motor activation might strengthen these connections needed in accessing phonological representations (Temple et al., 2003). Importantly, however, we also expect that auditory-motor activity associated with the speech training interacts with brain activity associated with gaming elements that have been suggested to enhance learning-induced plastic changes in the brain (Junttila et al., 2022; see also Oei and Patterson, 2013; Nahum and Bavelier, 2020, for game effects beyond language).

Producing words aloud in the game enables an automatic speech recognizer (Karhila et al., 2017, 2019) to assess the utterances and to give on-line feedback: player's performance is rewarded with one to five stars as a function of its accuracy. Players need to collect the stars to proceed in the game. This is expected to activate the reward system of the brain and to motivate players to produce foreign speech as accurately as possible. Together, these processes could typically result in the efficient establishment of accurate brain representations for foreign speech sounds and words. However, some dyslexic readers have been found to show atypical structure or function of striatum (Krishnan et al., 2016), which may affect the ability to learn from the game. Since striatum is the core area of the reward system of the brain, atypical function of striatum may hinder the processing of feedback that the game provides to reinforce learning. This may diminish the benefits of feedback and the game-based approach for learning. Therefore, we hypothesized that if the processing of feedback in striatum is indeed atypical in dyslexia (Krishnan et al., 2016), getting feedback from the game could benefit typical readers more than dyslexic readers and result in weaker game-based training effects in dyslexia.

To assess the effects of targeted training with the game and the non-game, in particular, long-term memory representations established in the brain, we used auditory event-related potential (ERP) measurements. Specifically, we focused on the mismatch negativity (MMN), an ERP component which is elicited when a stimulus violates a previously formed prediction (Winkler, 2007). The MMN can be elicited by presenting a regular sound stream consisting of so-called standard sounds and occasional deviant sounds (oddball paradigm). Its peak is typically observed 100–250 ms after the deviance onset and it is largest at fronto-central electrode sites when referenced to the average of mastoids (Sams et al., 1985). The MMN is considered particularly feasible for the present study for two reasons. Firstly, in addition to acoustic deviance the amplitude of MMN is enhanced by long-term memory traces for speech sounds (Näätänen et al., 1997) or individual spoken words (Pulvermüller et al., 2001). Processing ability reflected by the MMN governs attentive discrimination (Tiitinen et al., 1994; Amenedo and Escera, 2000), and previous training studies (Tremblay et al., 1998; Ylinen et al., 2010; Tamminen et al., 2015; Junttila et al., 2022) show that the MMN is feasible to tap the establishment of brain representations for non-native speech sounds and words that are in the focus of the present study. Therefore, measuring the MMN response allowed

us to evaluate plastic changes in brain due to behavioral foreign-language training. Secondly, the elicitation of MMN does not require attention (Näätänen et al., 1978; Fitzgerald and Todd, 2020), and thus it is a very useful tool for studying the target group of this study, namely, children who may have short attention span or who may not be motivated to perform tasks (Ylinen et al., 2021). Previous MMN studies have also robustly shown differences in speech processing between individuals with dyslexia and their controls (e.g., Schulte-Körne et al., 1998; Bruder et al., 2011; van Zuijen et al., 2013).

2. Materials and methods

2.1. Ethics

The study was conducted according to the Declaration of Helsinki and it was approved by the University of Helsinki Ethical Review Board in the Humanities and Social and Behavioral Sciences. Participation was voluntary. A written informed consent was acquired from the participant's guardians and an oral informed consent from the participants. The compensation for participation was one cinema ticket per hour of testing and EEG measurement.

2.2. Participants

The study included two groups of participants: children with dyslexia (N = 24) and fluently reading control children (N = 24). The control children are a subsample of the participants in a previous study (Junttila et al., 2022). The control children were selected so that they matched the children in dyslexia group by age and sex. According to parental reports, the participants were 7–11 years old monolingual native speakers of Finnish. The dyslexia group included 14 girls and 10 boys (mean age 9 years 8 months, SD 12 months). Correspondingly, the control group consisted of 14 girls and 10 boys (mean age 9 years 8 months, SD 10 months).

The participants were screened with pre-tests that measured literacy, cognitive skills, and phonological skills (see details below) to verify that they met the inclusion criteria and to decide the allocation to the two groups. The inclusion criteria for all participants of the study were as follows: age of 7–11 years, being a Finnish-speaking monolingual with no head injuries, having normal hearing, normal vision or vision corrected to normal with eyeglasses, and obtaining a minimum of 6 standard points in the four Wechsler Intelligent Scale for Children IV (WISC-IV; Wechsler, 2010) subtests used in the study (Block Design, Digit Span, Coding, and Vocabulary). There were also additional criteria that were different for the dyslexia group and the control group. To be included in the dyslexia group, the participants had to meet the following criteria: a maximum of 6 standard points in the Finnish reading and writing skill test LukiLasse (Häyrinen et al., 1999) word reading task and either a maximum of 6 standard points in the LukiLasse dictation task or a maximum of level score 3/9 in another Finnish literacy test task, Ala-asteen lukutesti (ALLU; Lindeman, 1998) word segmentation task. The additional inclusion criteria for the control group were a minimum of 7 standard points in the

LukiLasse word reading and dictation tasks and a minimum of level score 4 in the ALLU word segmentation task.

2.3. Cognitive pre-tests

Participants' knowledge of the target words used as stimuli during the EEG experiment was tested before participating in the first EEG measurement. Each stimulus word and pseudoword was presented *via* headphones one by one. After presenting each word, the children were asked whether the heard items sounded familiar to them. They were also asked whether they knew the meanings of the words that sounded familiar. The same task was conducted after training to indicate learning to recognize the words and learning the meanings of the words.

Participants' reading, writing, cognitive, and phonological skills were assessed with standardized neuropsychological tests in Finnish. Writing skills were assessed with LukiLasse dictation task where participants are required to write dictated words and sentences without time limit. Reading skills were assessed with LukiLasse word reading task and ALLU word segmentation task. LukiLasse word reading task is used to assess reading accuracy and speed. It requires reading aloud a list of words within a time limit of 2 min. In a similar vein, ALLU word segmentation task assesses technical reading skills. In this task, children read strings of letters without spaces and segment the strings into words by drawing vertical lines between the word boundaries within a time limit of 3.5 min. The test has two standardized versions (A and B) with a different set of word strings. The children did the version A of the ALLU test during the pre-test session before the training period and the version B after the training period right before the final EEG recording. In addition to literacy skills, other cognitive skills that may affect performance in our tasks were measured to ensure that the two groups did not differ in this respect. Perceptual reasoning skills were assessed with block design subtest from the Wechsler WISC-IV, auditory short-term memory was assessed with WISC-IV forward and backward digit span, eye hand coordination and processing speed were assessed with WISC-IV coding, and vocabulary was assessed with WISC-IV vocabulary.

In addition, since dyslexia is typically characterized by slow naming and poor phonological awareness, Rapid Automatized Naming Test (RAN, Ahonen et al., 2003) subtests colors and letters were used to assess naming skills and Common Unit test was used to assess phonological awareness. In Common Unit test the child is presented with spoken pseudoword pairs *via* headphones. The task is to name the common phoneme in the two words. First, there is a practice phase where three practice pairs are presented. The experimenter corrects the wrong answers and tells what the right one is in the practise phase. After that, there are 15 trials where each pair is presented once. The length of the pseudowords vary between 5 and 12 Finnish phonemes. Each correct answer scores one point and no feedback is given during the test trials.

The test performance was compared between the groups with independent samples *t*-tests (Table 1). No differences between the groups were found in WISC-IV block design, WISC-IV vocabulary, WISC-IV coding, and RAN colors. In addition to tests used for group allocation (LukiLasse word reading, LukiLasse dictation, ALLU word segmentation), groups also differed in WISC-IV digit

TABLE 1 Participants' literacy, cognitive, and phonological skill scores.

Test	Children with dyslexia: standard score/skill level/raw score (±SD)	Controls: standard score/skill level/raw score (±SD)	t-test
LukiLasse word reading	3.29 (±1.83)	10.13 (±2.42)	$t = -11.04$, $p = <0.001$
LukiLasse dictation	5.13 (±3.44)	10.54 (±1.35)	$t = -7.18$, $p = <0.001$
ALLU word segmentation	1.88 (±0.68)	4.71 (±1.04)	$t = -11.16$, $p = <0.001$
WISC-IV block design	10.17 (±3.32)	10.42 (±2.81)	$t = -0.28$, $p = 0.780$
WISC-IV digit span	8.46 (±1.53)	10.58 (±3.27)	$t = -2.88$, $p = 0.006$
WISC-IV vocabulary	10.58 (±2.47)	10.79 (±1.61)	$t = -0.35$, $p = 0.731$
WISC-IV coding	10.13 (±3.01)	10.75 (±3.04)	$t = -0.72$, $p = 0.478$
RAN colors	53.96 (±10.25)	49.79 (±13.89)	$t = 1.18$, $p = 0.243$
RAN letters	38.13 (±9.17)	30.42 (±5.18)	$t = 3.59$, $p = <0.001$
Common unit	4.33 (±3.32)	9.92 (±2.69)	$t = -6.41$, $p = <0.001$

Group means are given as standard scores for LukiLasse and WISC-IV, skill level for ALLU, and raw scores for RAN (time in seconds) and common unit.

span, RAN letters, and Common Unit task, where the control group performed better than the children with dyslexia. These differences were expected because impairments in these skills are typical in dyslexia (Wagner and Torgesen, 1987; Vellutino et al., 2004; Peters et al., 2020; Andreola et al., 2021).

2.4. EEG experiment

2.4.1. EEG stimuli

Similarly as in our previous gaming study (Junttila et al., 2022), the stimuli presented during the EEG recordings were spoken English words and pseudowords with voiceless and voiced dental fricative phonemes /θ/ and /ð/, respectively. These phonemes were chosen because they are not part of the Finnish phonological system, and we wished to ensure that the stimulus words contain phonemes foreign to Finnish speakers to study learning effects. Although the children may have been exposed to these phonemes before the experiments, they had not learned to pronounce or distinguish the phonemes. The choice of the phonemes is also supported by the fact that syllables containing /θ/ as a foreign phoneme have previously been shown to elicit an MMN response in a similar age group (Jost et al., 2015). In addition, we aimed to choose words that would be unfamiliar so that the participants would not know their meanings in advance. Therefore, we checked English textbooks commonly used in schools to choose words that were likely not included in their vocabulary. This way we ended up choosing the words *healthy* [ˈhelθi] and *feather* [ˈfeðə] and formed corresponding pseudowords from them by replacing the target fricatives with stops (resulting in *healty**

[ˈhelti]* and *feder* [ˈfedə]*). The stimuli were recorded in a sound-attenuated studio. A list of English words and their minimal-pair counterpart pseudowords were given to a native English speaker who pronounced them several times. The best exemplars of the words *healthy* and *feather* and pseudowords *healty** and *feder** were chosen from the recordings to be used as stimuli.

The recorded words and pseudowords were modified using Praat software Version 5.1.45 (Boersma and Weenink, 2010). To make sure that the word and the pseudoword could not be distinguished before the target phonemes, we used beginnings of the words *healthy* and *feather* as the beginnings of the pseudowords (Figure 1). The word *healthy* was cross-spliced at zero-crossing at 200 ms and the first part of the word was used as the beginning of the pseudoword *healty**. Similarly, the word *feather* was cross-spliced at zero-crossing at 220 ms and the first part was used as the beginning of the pseudoword *feder**. The pitch contours of *healthy* and *feather* were adopted to *healty** and *feder**, respectively.

As the aim of the study was to compare within-subject effects of the game training and the non-game training regardless of the foreign phoneme type, the responses to both *feather* and *healthy* were combined for analysis purposes.

2.4.2. EEG recordings

EEG was measured on two or three occasions, namely, a baseline measurement, a pre-training measurement, and a post-training measurement. All children participated in the pre-training measurement and the post-training measurement, whereas only part of the children participated in the baseline measurement (20 of the children in the dyslexia group and 16 of the children in the control group). The baseline measurement took place prior to the pre-training measurement, and it was conducted to assess and control for the effects of repeated EEG measurements and out-of-game exposure to English (nevertheless, inability to participate in three measurements was not considered an exclusion criterion). The time between the baseline measurement and the pre-measurement, as well as between the pre-measurement and the post-measurement was 5 weeks on average (Figure 2). The participants played the language learning game containing both the game levels and the non-game levels for training the target phonemes between the pre-measurement and the post-measurement.

EEG was recorded with BioSemi ActiveTwo system (BioSemi Inc., Amsterdam, and The Netherlands) and with Biosemi ActiView Version 7.07 EEG acquisition software using a 64-channel cap. Additional electrodes were placed on the tip of the nose and the mastoids for re-referencing, near the outer canthi of the eyes for horizontal eye movements, and under the left eye for vertical eye movements. The measurements were conducted in a sound and electrically attenuated room, where the participants sat on a comfortable chair. The participants chose a film to watch during the measurement. The film was muted but the subtitles were shown in Finnish. Headphones were used to present the auditory stimuli binaurally. The volume was set at a comfortable hearing level of 60 dB. The participants were instructed to ignore the stimuli and to concentrate on the film instead.

The stimuli were presented in four different sequences: two oddball sequences and two standard-only sequences. The oddball sequences had a repeating standard pseudoword that was occasionally replaced by a deviant word. A standard pseudoword
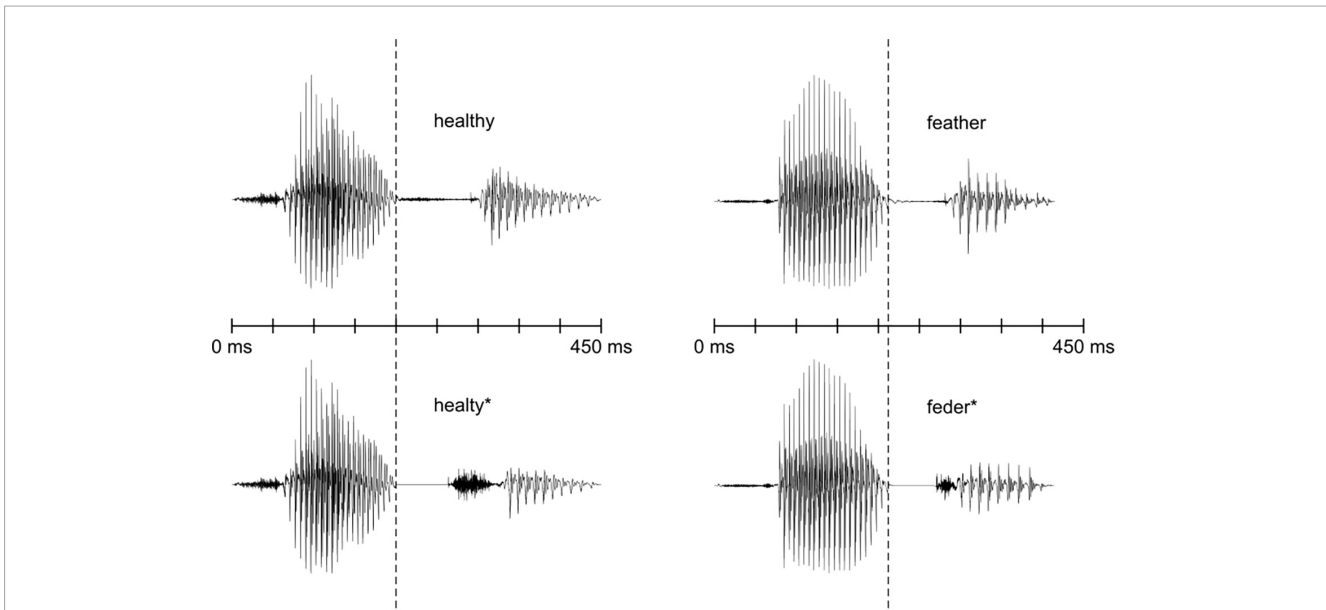
FIGURE 1
Waveforms of the stimulus sounds. The time points where the pseudowords deviate from their word counterparts are marked with dashed lines.
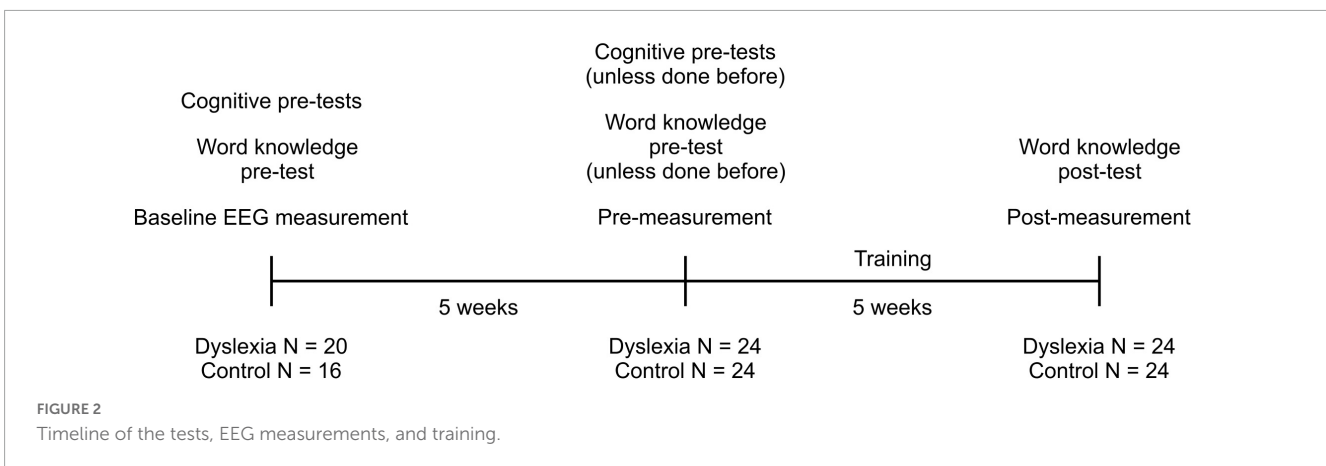

FIGURE 2
Timeline of the tests, EEG measurements, and training.

feder* ($p = 0.8$) and a deviant word *feather* ($p = 0.2$) were used in one oddball sequence focusing on /ð/ and a standard pseudoword *healty** ($p = 0.8$) and a deviant word *healthy* ($p = 0.2$) were used in the other one focusing on /θ/. In both oddball sequences, the standard pseudoword was presented 480 times and the deviant word was presented 120 times. Standard-only sequences had the real words from the oddball sequences presented 180 times as repeating standards. The word *feather* ($p = 1$) was presented in one standard-only sequence and the word *healthy* ($p = 1$) in the other. All sequences had an inter-stimulus interval (ISI) of 500 ms. Presentation® software (Version 17.2, Neurobehavioral Systems, Inc., Berkeley, CA)[1] was used to present the stimulus sequences. The four sequences were presented in a pseudorandom order for each participant and measurement session so that the presentation order of sequences containing words trained with the game and words trained with the non-game were counterbalanced. Each EEG measurement session lasted 2 h, including preparation.

---

1   www.neurobs.com

2.4.3. EEG data processing

The EEG data processing was conducted with BESA Research 7.0 software (BESA GmbH, Gräfelfing, Germany). Bad EEG channels were interpolated using the data from other channels. To improve the signal-to-noise ratio, the data were re-referenced to the average of left and right mastoids. Children's EEG data are typically noisier than those of adults. Thus, in line with our previous work (Junttila et al., 2022), a band pass filter of 1.5–20 Hz with a slope of 24 dB/octave was used to remove slow drifts from the data that are beyond the MMN range and could distort the MMN amplitude quantification (Picton et al., 2000; Sinkkonen and Tervaniemi, 2000; Nenonen et al., 2005; Kalyakin et al., 2007; Jost et al., 2015). BESA Research Artifact Correction was used to detect and correct the horizontal and vertical eye movements. The data were then divided into epochs using a −100–800 ms time window relative to the stimulus onset. The 5 first epochs of each sequence were rejected, as were all epochs contaminated by artifacts exceeding ± 75 µV. Average waveforms were created for each stimulus type separately using the accepted epochs. Responses

to the stimuli derived from standard-only blocks were subtracted from the responses to identical stimuli in deviant positions to form difference waveforms, and the baselines were corrected at the −100–0 ms pre-stimulus interval.

Grand average waveforms for the game condition and the non-game condition were formed for both groups separately. To compare the effects of the game and the non-game, the ERP waveforms for the game condition and the non-game condition were constructed by combining the responses of both target words. The MMN response is typically most prominent on fronto-central regions in data referenced to mastoids, and thus nine electrodes (F3, Fz, F4, FC3, FCz, FC4, C3, Cz, and C4) from that area were selected for analysis. The peak latency of the MMN response was determined from a grand-average waveform over both groups, both conditions, all time points, and all nine selected electrodes. The mean MMN amplitudes for both groups, both conditions, all time points, and all electrodes were then quantified from individual difference waveforms (deviant minus standard-only, i.e., the responses to the same stimulus) using a ± 20 ms time window centered at this average peak latency of 334 ms.

2.5. Training

Between the last two EEG measurement times, the participants played "Say it again, kid" (SIAK) that is a language-learning game targeting spoken English and running on Windows laptops or Android tablets. The players used microphone headsets for playing. In the SIAK game, the players can move around on a game board and whenever they encounter a card, the card pops up and introduces an English word produced by a native English speaker (**Figure 3**, left). The card shows a picture associated with the word and the players hear the word spoken in Finnish to ensure they understand the meaning of the word. Then they hear the word spoken in English. The players' task is then to say the word aloud in English, trying to imitate the model pronunciation as closely as they can. The players' utterances are recorded and played back to them along with the original English pronunciation. An automatic speech recognition (ASR) technology, which is optimized for children's voices (Karhila et al., 2017, 2019), evaluates the players' utterances and, as feedback, awards one to five stars based on how close it was to the model pronunciation. The total number of stars the player has scored is shown on the top right-hand corner of the game screen (**Figure 3**, left). Scoring more stars opens new paths on the game boards. The game has visually different levels, and each level ends with a test card, where players can test their learning. The test card shows a picture of one word learned during the level but does not play a model pronunciation. The players need to produce the correct word and score at least three stars to unlock a next level. If the players at first do not succeed, they can replay the card containing the test word, practice pronouncing the word as many times as they like, and then go back to the test card to try to score enough stars.

The children played the game on average 16.3 min (control 16.8 min, dyslexic 15.7 min) per play session. They played on average 1.1 play sessions (control 1.0, dyslexic 1.1) a day on 2.8 days (control 2.7, dyslexic 3.0) a week during a period of 4.3 weeks (control 4.4, dyslexic 4.2). Altogether, the children with dyslexia

played the game 209 min (SD = 42.9, range 149–323 min) and the control children played the game 197 min (SD = 48.9, range 145–341 min). All children completed all levels of the game. An independent samples t-test was conducted to compare the playing times between the control group and the dyslexia group. No significant difference was found [$t(46)$ = 0.87, p = 0.39, 95% CI = −15.17, 38.26].

The SIAK game had altogether 27 levels, yet only six of them contained words with the target phonemes /θ/ (e.g., *healthy, mouth, three*) and /ð/ (e.g., *feather, mother, this*). Out of these six levels, three were normal game levels with the target sounds and three were non-game levels with the target sounds. The non-game levels had English word imitation tasks for the target sounds similarly to the game levels but did not have any game-like elements (**Figure 3**, right). In the non-game levels, the players were not free to explore and choose the order of the items like in the game levels, but rather the presentation order of the words was forced. Instead of the colorful game board of the game levels, the non-game levels had a black arrow on a white background. Also, the players did not score stars or receive other feedback in the non-game levels. Otherwise, the non-game training was similar to the game training: Players heard spoken words in Finnish and English and they needed to imitate the English ones. The English stimuli were counterbalanced across the level type: Half of the control group and half of the dyslexia group trained words containing the phoneme /θ/ with the game levels and words containing /ð/ with the non-game levels. The other halves of the control and dyslexia groups trained words containing the phoneme /ð/ with the game levels and words containing /θ/ with the non-game levels.

There were about 15 different words to learn at each level of the game. The players could freely explore the game levels and decide how many times they practised each word (with the exception of the non-game levels where the order and number of words were fixed). The number of times words were practised with the game levels and with the non-game levels was controlled within participants. To ensure equal amount of practice of the words containing the target phonemes at the game levels (levels 16, 20, and 25) and the non-game levels (levels 17, 21, and 26), the free-to-explore game levels with the target sounds were always presented before the non-game levels with the other target sounds. The number of times players practiced words with the target phonemes and EEG stimulus words at the game levels was recorded. During the following non-game level, words with the other target phonemes as well as the EEG stimulus words were presented a matching number of times. Thus, each player got equal amount of training consisting of listening to and producing speech with the game and the non-game.

Word learning during training, as indicated by familiarity with the word form and knowledge of word meaning, was tested after the training period just before the EEG post-measurement similarly as before the first EEG measurement in both groups. In addition, to assess the learning of sound production for two target phonemes /θ/ and /ð/ by children with dyslexia, we evaluated their utterances collected by SIAK. The utterances were evaluated using the automatic speech recognition algorithm (Karhila et al., 2019) used in the game. We counted the relative number of utterances containing the correct /θ/ and /ð/ phonemes out of all attempts for the first session and the last session containing these phonemes. The learning was defined as a change between the relative numbers of correct utterances.
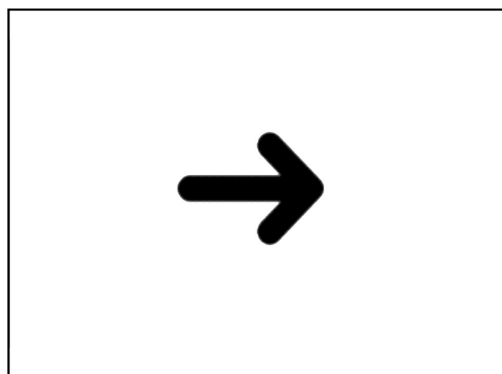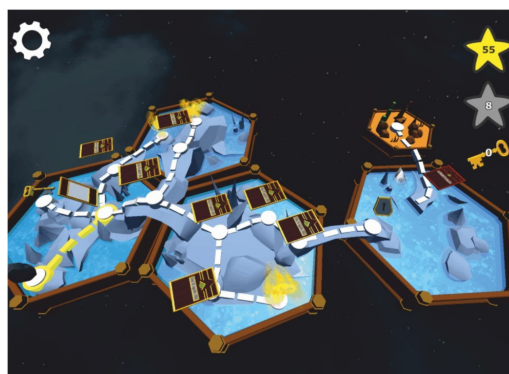
FIGURE 3
An example of a game level **(left)** and a non-game level **(right)**.

2.6. Statistical analyses

Differences in the learning of words, trained with the game or with the non-game (i.e., the same words as those used as EEG stimuli), were examined with McNemar's tests. Two different types of learning were assessed: learning to recognize the word form (familiarity) and learning the meaning of the word. The tests were conducted for both groups separately. In addition, we used paired one-tailed t-tests to examine the learning to produce the target phonemes correctly in the dyslexia group.

Differences in MMN amplitude change between the two groups after training with the game or the non-game (post-test minus pre-test) were compared with a linear mixed model analysis. The analysis was conducted with IBM SPSS Statistics, Version 27 software. The analysis included Group (dyslexic – fluent), Treatment (game – non-game), Anteriority, and Laterality as factors and the change in MMN amplitude between the baseline measurement and pre-training measurement as a covariate. A separate paired-samples t-test was conducted to compare

difference between baseline and pre-test within the dyslexia group's game condition.

To study the connections between neural training effects and literacy, RAN, and phonological awareness, we performed correlation analyses (Pearson's r). To this end, the MMN amplitude change after training with the game or the non-game was averaged across the nine fronto-central electrodes. Literacy score was calculated by taking the average of the z-scores of LukiLasse word reading task, LukiLasse dictation task, and ALLU word segmentation task. RAN score was the average of RAN speed z-scores of color and letter tasks. In both composite scores, negative values indicate below-average and positive values above-average performance. Benjamini-Hochberg correction (Benjamini and Hochberg, 1995) was used to adjust for multiple comparisons. False discovery rate (FDR) was set at 0.05.

3. Results

Both EEG stimulus words sounded familiar to one child in the control group and four children in the dyslexia group before the first EEG recording session. In addition, 11 children in the control group and 8 children in the dyslexia group found one of the words familiar sounding. None of the children in either group knew what the words meant. After the training period, the stimulus words learned with the game sounded familiar to 20 children in the control group and 19 children in the dyslexia group. The stimulus words learned with the non-game sounded familiar to 17 children in the control group and 22 children in the dyslexia group. No familiarity difference between the words learned with the game or the non-game was found in either group according to the McNemar's test (control: $p = 0.453$, dyslexia: $p = 0.453$). After training with the game, 12 children in the control group and 9 children in the dyslexia group had learned the meanings of the words, whereas after training with the non-game, 9 children in the control group and 5 children in the dyslexia group had learned the meanings. According to the McNemar's test, no difference between the meanings learned with the game and the non-game was, however, found in either group (control: $p = 0.549$, dyslexia: $p = 0.219$). For speech production, t-tests revealed a difference for the children with dyslexia in learning to produce the
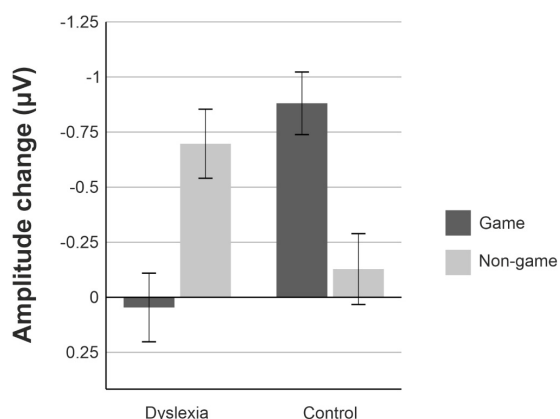

FIGURE 4
The mean amplitude changes between the pre-training and post-training measurements for the dyslexia group and the control group. Amplitude changes are averaged over the nine electrodes selected for analysis. Error bars represent the standard error of the mean.

correct phonemes with the game and the non-game [$t(22) = 1.78$, $p = 0.044$, 95% CI = -0.01, 0.16]. The learning was on average 7.5 percentage points stronger in the non-game situation than in the game situation.

For the MMN data, the linear mixed model revealed a significant interaction of Group and Treatment, [$F(1, 550.53) = 17.19$, $p = <0.001$] (**Figures 4**, **5**). The estimated marginal means were evaluated at the average MMN amplitude change between the baseline measurement and pre-measurement (-0.31 μV) (**Table 2**). Pairwise comparison showed that in the dyslexic reader group, training with the non-game increased the MMN amplitude more (-0.66 μV) than training with the game (-0.07 μV) ($p = 0.001$, 95% CI = 0.23, 0.95) (**Figures 5**, **6**). In the control group, training with the game increased the MMN amplitude more (-0.88 μV) than training with the non-game (-0.38 μV) ($p = 0.013$, 95% CI = 0.10, 0.89). No significant difference was found between baseline and pre-test within the dyslexia group's game condition [$t(19) = 1.48$, $p = 0.156$, 95% CI = -0.24, 1.42].

Correlation analyses revealed a significant correlation between phonological awareness and MMN amplitude change after training with the non-game in the dyslexia group [$r(22) = 0.50$, $p = 0.012$]. A larger MMN amplitude increase (more negative) was associated with poorer phonological awareness (**Figure 7** and **Table 3**). No other statistically significant correlations were found (**Tables 3–6**).

4. Discussion

This study aimed to investigate whether playing a digital language-learning game or training with a non-game would benefit the foreign-language learning of children with dyslexia and induce more plastic changes in the brain as indicated by MMN responses. We compared the change in MMN amplitude before and after training with the game and the non-game in children with dyslexia and typically reading control children. In addition, we examined whether articulatory or phonological learning with the game or the non-game is connected with literacy, rapid naming, and phonological awareness, which are typically impaired in dyslexia. In the dyslexia group, training with the non-game increased the MMN responses more than training with the game. By contrast, in the control group, training with the game was more effective than training with the non-game. In the dyslexia group, the MMN amplitude increased more after training with the non-game in those children who had poorer phonological awareness. The children with dyslexia also learned to produce the target phonemes better with the non-game than with the game.

The children with dyslexia showed larger change in the MMN responses after training with the non-game than with the game and thus appeared to benefit more from training with the non-game. The control group showed an opposite effect, benefiting more from training with the game. Because this finding is based on an interaction effect between Group and Treatment, it cannot be explained, for example, by acoustic properties of the stimuli. The pattern of results cannot be explained by other stimulus properties either, because the words and phonemes learned with the game and with the non-game were counterbalanced, so that half of the children within each group learned the target word *healthy* with

the game and the target word *feather* with the non-game and vice versa. The ERP waveforms for the game and non-game conditions were created for each group separately by combining the responses to *healthy* and *feather* trained with the game and the responses to *healthy* and *feather* trained with the non-game, respectively. Thus, differences in word or phoneme difficulty or frequency cannot account for the effect. Further, each child trained words both with the game and with the non-game. Therefore, learning differences between the game and the non-game were compared in the same children within the dyslexia group, and likewise within the control group. Accordingly, training effects between the game and the non-game condition within the groups cannot be explained by individual differences between children training with the game and the non-game. Note also that the interaction between Group and Treatment was significant after controlling for the effects of repeated EEG measurements or other exposure with a baseline measurement that was used as a covariate in statistical tests. It is, therefore, unlikely that our training effects reflected in the interaction were explained by spontaneous learning during the EEG measurements or constant exposure to English outside the training (e.g., in classroom, *via* media). The different gains in the MMN amplitude in dyslexia group and the control group are, therefore, best explained by the type of training, that is, whether the training of the target sounds was conducted with the game or the non-game.

Although we have tried to control for factors that could affect our training effects, the data pattern may, nevertheless, rise some questions. Looking at MMN waveforms and scalp topographies, one may be puzzled by the difference between baseline and pre-test MMNs in the dyslexia gaming condition: compared to the baseline, pre-test MMN seems enhanced. To determine the consistency of this difference, we separately compared these conditions within this group with a paired-samples *t*-test. The difference was not significant, suggesting that it was not consistent. Perhaps the fact that not all children participated in the baseline measurement accounts for some variation in the amplitude. One may also ask whether watching subtitled films during the MMN measurements may have affected the participant groups differently. Using subtitled films is a standard procedure in MMN studies, including those addressing spoken words (Pulvermüller and Shtyrov, 2006). No effect of subtitles on MMN has been found during speech stimulation (Pettigrew et al., 2004) and subtitled films have been used also in children with dyslexia (Widmann et al., 2012). We cannot fully rule out the possibility that subtitles could affect the two groups differently, although even if this was the case, the effects are more likely to affect obligatory responses than the MMN (Pettigrew et al., 2004). However, it is very unlikely that the use of subtitles accounts for the current effect since our results are based on an interaction rather than a group main effect. The gaming effect cannot be directly caused by the subtitles because the gaming manipulation was present only in training and not in the MMN measurement. Any stimulus effects interacting with the subtitles are ruled out by counterbalancing the stimuli across training types. The means by which subtitles could potentially modify the interaction indirectly is by affecting children's attention differently, if that had an effect on MMN. However, a thorough review by Sussman (2007) states that although attention may affect MMN elicitation through task requirements and regularity formation as part of auditory scene analysis, simply attending to deviants does not alter the
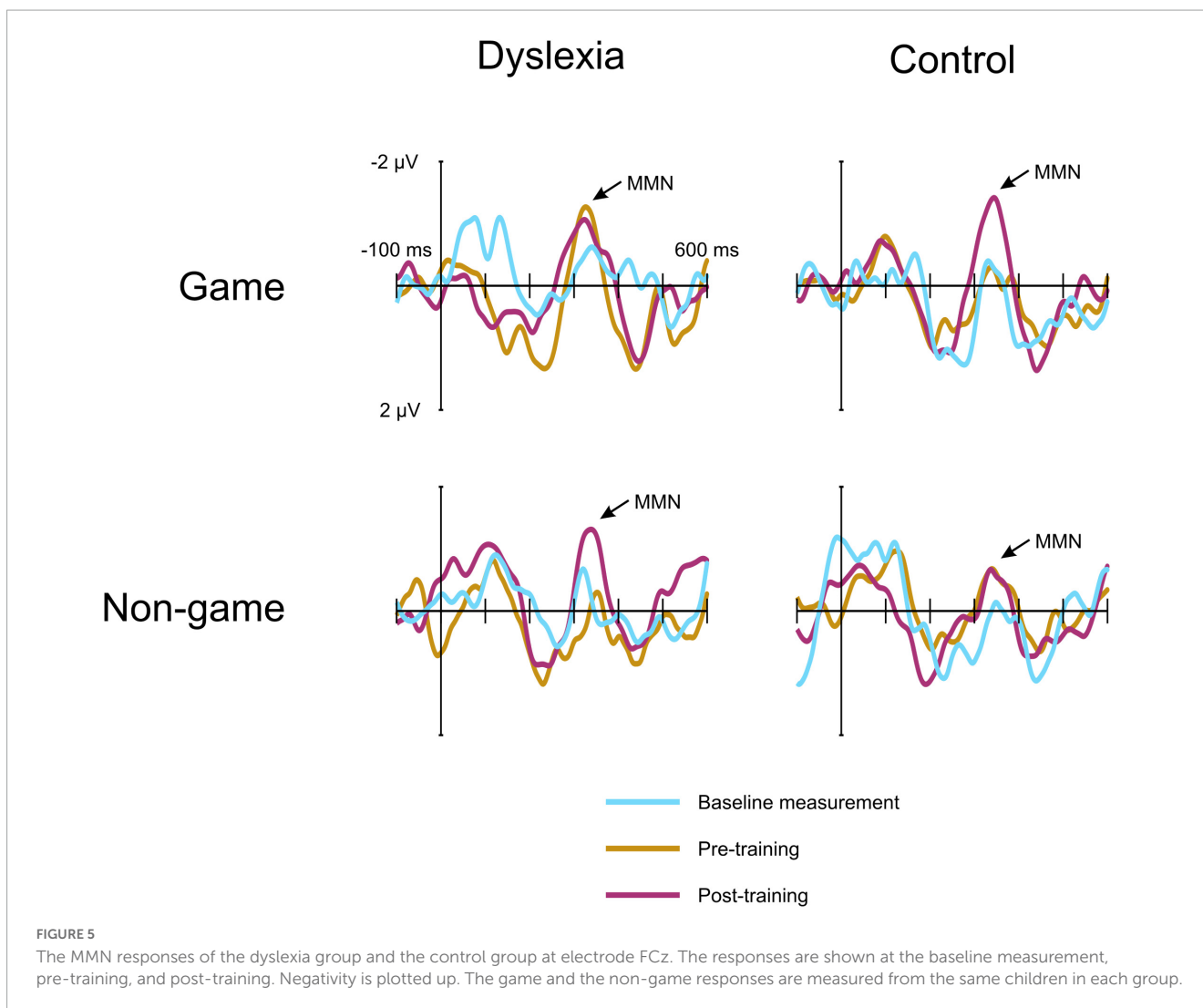
FIGURE 5
The MMN responses of the dyslexia group and the control group at electrode FCz. The responses are shown at the baseline measurement, pre-training, and post-training. Negativity is plotted up. The game and the non-game responses are measured from the same children in each group.

amplitude of the MMN component. This suggests that attention effects do not explain the interaction. Therefore, we argue that the interaction is due to activation of brain representations that has been differently influenced by the training type in the two groups.

The difference between the current game and the non-game lies in feedback, freedom of choice, and rich visual features in the former, as opposed none of them in the latter. All these features may

TABLE 2 Estimated marginal means (EMM) of the MMN amplitude changes between the pre-measurement and the post-measurement in µV for the dyslexia group and the control group for the game and non-game conditions.

| Group | Condition | EMM | S.E. | 95% CI | |
				Lower	Upper
Dyslexia	Game	−0.07	0.32	−0.70	0.57
	Non-game	−0.66	0.32	−1.30	−0.01
Control	Game	−0.88	0.36	−1.61	−0.15
	Non-game	−0.38	0.36	−1.12	0.35

The estimates are evaluated at the average MMN amplitude change between the baseline measurement and the pre-measurement.
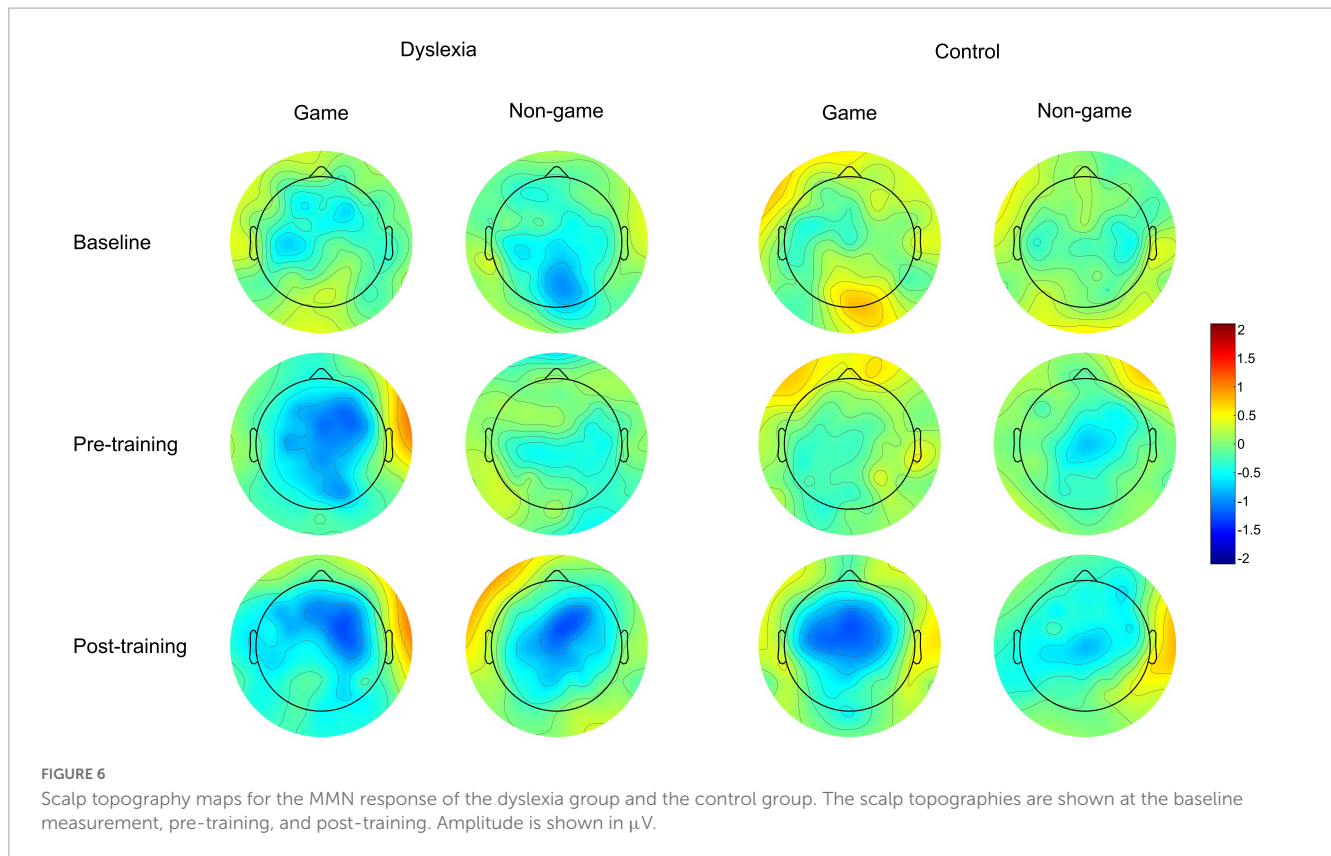
activate the reward system of the brain (see Marco-Pallarés et al., 2007; Wittmann et al., 2007; Leotti and Delgado, 2011; Murayama et al., 2015). In the game, the players gained stars as feedback based on how well they imitated the model word. In the non-game, the players did not get any feedback. Differences in the function of the reward system of the brain may explain why children with dyslexia benefited more from the non-game training, whereas the control children benefited more from the game training. In typical learners, feedback during speech and speech-sound learning activates striatal structures and this activation contributes to successful learning (Yi et al., 2016; Feng et al., 2019). Children in the control group benefited more from playing the game than training with the non-game, which is likely due to the activation of the reward system in the brain. Since it has previously been argued that dyslexia is characterized by atypicality in the striatum (Krishnan et al., 2016), in the dyslexia group the inefficiency of striatal functioning may result in poorer learning of words and speech-sounds in the game. Thus, it is plausible that the dyslexia group did not benefit from the feedback, freedom of choice, or novel visual features provided by the game the way the control group did.

Unlike the non-game, the game condition included colorful graphics in form of a game board with multiple choices of
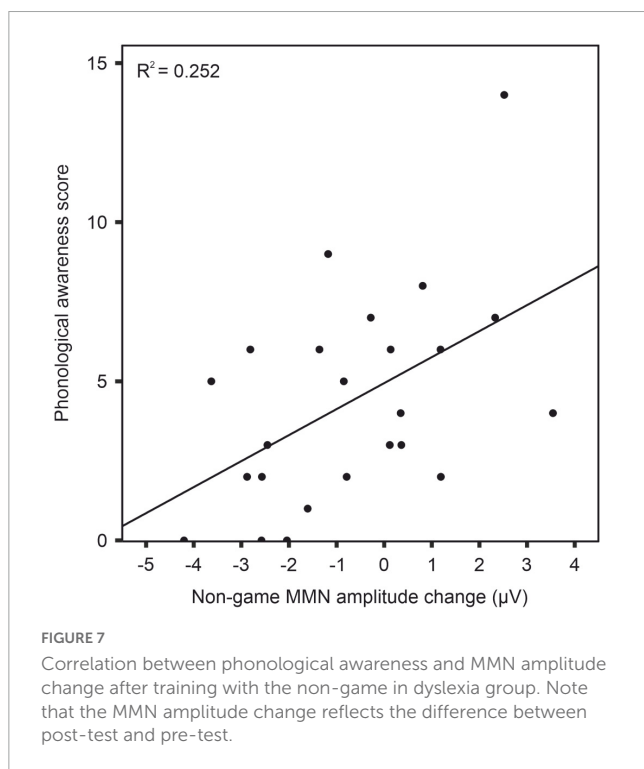
FIGURE 6
Scalp topography maps for the MMN response of the dyslexia group and the control group. The scalp topographies are shown at the baseline measurement, pre-training, and post-training. Amplitude is shown in μV.

what to explore next. These features may increase, for example, children's level of arousal, interest, engagement, and attention (Park and Lim, 2007; Oei and Patterson, 2013). These processes have likely enabled gaming elements to enhance the learning of the

FIGURE 7
Correlation between phonological awareness and MMN amplitude change after training with the non-game in dyslexia group. Note that the MMN amplitude change reflects the difference between post-test and pre-test.

control children, which is in line with previous studies on gaming benefits for children's foreign-language learning (Liu and Chu, 2010; Aghlara and Tamjid, 2011; Franciosi, 2017; Tsai and Tsai, 2018). A previous study found benefits of hybrid technology use on foreign language receptive vocabulary learning (translation and writing) for children with dyslexia (Eden and Shmila, 2022). That study compared a hybrid technology intervention to a traditional technology intervention, where both interventions had similar visual elements and provided feedback to the learners. In the present study, however, the children with dyslexia benefited more from the non-game training, which was stripped of all the game-like elements and was thus visually much simpler. Furthermore, no feedback was given in the non-game training. Because of these differences in the study design, the findings of the present study are not in contradiction with the findings of Eden and Shmila (2022). Rather, the differences might shed light on why the non-game training was more effective for children with dyslexia. Since children with reading difficulties have been found to also have difficulties in inhibiting unattended visual information (Facoetti et al., 2003), the lack of rich visual elements may have been one factor contributing to the finding that the children with dyslexia benefited more from the non-game training (however, see **Figure 7** for individual variation within the dyslexia group in the non-game). Thus, at least in those individuals with more learning benefit from the non-game than from the game, attention may have been drawn to other elements in the gaming view than to just the current card and the word to learn. Another possibility may lie in the individual differences in optimal arousal level for learning (Hallam et al., 2002). The arousal level induced by the game may have been too high for some individuals with dyslexia, although for some others

TABLE 3 Correlations in the dyslexia group between the test performance and MMN amplitude change after training with the non-game.

Test	r	Raw p-value	Rank	Critical value at 0.05
Phonological awareness	0.50	0.012	1	0.017
Literacy	0.17	0.439	2	0.033
RAN	−0.16	0.462	3	0.050

TABLE 4 Correlations in the control group between the test performance and MMN amplitude change after training with the non-game.

Test	r	Raw p-value	Rank	Critical value at 0.05
RAN	0.16	0.445	1	0.017
Phonological awareness	−0.16	0.448	2	0.033
Literacy	−0.02	0.917	3	0.050

TABLE 5 Correlations in the dyslexia group between the test performance and MMN amplitude change after training with the game.

Test	r	Raw p-value	Rank	Critical value at 0.05
RAN	−0.24	0.260	1	0.017
Literacy	0.13	0.544	2	0.033
Phonological awareness	0.09	0.686	3	0.050

TABLE 6 Correlations in the control group between the test performance and MMN amplitude change after training with the game.

Test	r	Raw p-value	Rank	Critical value at 0.05
Literacy	0.47	0.022	1	0.017
Phonological awareness	0.38	0.068	2	0.033
RAN	−0.17	0.430	3	0.050

it may have been optimal. However, it is somewhat puzzling that earlier gaming studies have not shown such effects in children with dyslexia; rather, gaming has been shown to produce significant learning gains, for example, in word reading (Ronimus et al., 2019; Carvalhais et al., 2020). Perhaps then, gaming elements do not draw attention from the exposure or processes that are most critical for performing the task (e.g., here encoding of the phonological units is critical for their production, which is the main task). In contrast, the gaming elements could diminish the allocation of attention to exposure or processes that support learning but are not critical for the task at that particular moment. For example, if players wanted to proceed quickly on the game board in the current setup, they may have not focused on monitoring their speech or listening to their own utterances that were played to them after they had uttered each word, and they also may have not compared them to the model pronunciations that were also re-played. Thus, the game features may have distracted speech monitoring and self-assessment of one's utterances if they urged the child to proceed quickly in the game.

In the dyslexia group, learning with the non-game was correlated with phonological awareness. The children with poorer phonological awareness showed a larger increase in the MMN amplitude, which may reflect alleviation of their phonological processing difficulties. The non-game training thus benefited more those children who had weaker phonological skills to start with. The training, based on listening to and producing foreign language speech without possibly distracting game elements, may have specifically tapped phonological skills,

which has been found to benefit children with dyslexia (Snowling and Hulme, 2011). The connection between poor phonological awareness and learning with the non-game suggests that this kind of training may be especially helpful for the children with inferior phonological skills. No correlation was, however, found between the MMN amplitude increase and literacy scores, although dyslexia has been shown to impair learning foreign languages (Downey et al., 2000; Helland and Kaasa, 2005; Soroli et al., 2010; Ylinen et al., 2019). A possible explanation could be that the training period in this study was too short and literacy improvements in dyslexia require longer-term learning.

This study compared children with dyslexia and control children that were matched by age and sex. As children with dyslexia have poorer reading skills than the control children, comparing them with control children with matching reading-skill or phonological-skill level regardless of age could provide additional information. Therefore, further study is needed to investigate whether there are differences in the effectiveness of digital game-based foreign-language learning between children with dyslexia and reading-skill or phonological-skill matched control children.

Although significant learning effects were observed in brain responses, no significant differences between training with the game or the non-game were found in word familiarity or learning the meanings of the words in either group of children. As discussed in our previous paper (Junttila et al., 2022), brain responses may more strongly reflect foreign speech-sound learning, whereas word familiarity and learning the meanings may reflect word learning. Nevertheless, word familiarity results point to the same direction as the brain-response data: the dyslexia group tended to improve more in the non-game, whereas the controls tended to improve more in the game. In a similar vein, these and brain-response results are supported by the dyslexic children's speech production data, showing significantly larger improvement for the speech sounds trained in the non-game than in the game. The finding of similar patterns in MMN, speech production, and perhaps also word familiarity, but a different pattern with word meanings, supports our earlier interpretation that the learning effects induced by our paradigm concern primarily phonology and speech-sound processing rather than word meanings.

The results of this study provide information that can be used for developing foreign-language teaching practices and designing digital foreign-language learning applications. Whereas typically reading children benefit from game-based learning, training with an application that has a relatively simple layout and less possibly distracting elements may be more useful for children

with dyslexia – at least those with poorest phonological skills. As neural level individual differences can have an important role in language learning (Turker et al., 2021); it is noteworthy that we also saw large individual differences within the groups. Although at the group level we found a gain in the dyslexia group's MMN amplitude in response to training with the non-game, in some individuals the MMN amplitude decreased. It has been hypothesized that dopamine-related genes are associates with language-learning differences (Wong et al., 2012), which could shed light why elements activating the reward system support the learning of some individuals but not others. Therefore, offering different kinds of teaching and learning approaches and personalizing language learning could aid a wider group of learners.

Note that although in the present setup our non-game training increased the brain responses more than the game training in children with dyslexia, the current data does not, however, suggest that gaming does not work as a learning technique in dyslexia. For example, gaming could motivate children with dyslexia to rehearse useful skills longer than a non-game application, and this kind of long-term practice could bring the children the intended benefits, whereas a non-game application that is eventually not used at all would not. The present results also do not tease apart the effects of different game features. Thus, the differences found here in learning with the game and the non-game may be related to one or two features of the game or a combination of all features, and their effect may also be individual. Further investigation of training techniques optimal in dyslexia is, therefore, needed. Nevertheless, we may speculate that visually simple design might be beneficial for many children with dyslexia.

To conclude, the results show that unlike typically reading children, children with dyslexia may benefit more from visually simple training than from visually rich game training. In the dyslexia group, the simple non-game training enhanced the MMN amplitude more than the game training. The non-game training increase in the dyslexia group was connected to phonological awareness evaluated before the training: the MMN increase was larger in the children with poorer phonological awareness. This could indicate that articulatory training with a simple training application could remediate some spoken foreign-language learning difficulties related to phonological processing deficits.

Data availability statement

The raw data supporting the conclusions of this article will be made available by the authors, without undue reservation.

Ethics statement

The studies involving human participants were reviewed and approved by the University of Helsinki Ethical Review Board in the Humanities and Social and Behavioral Sciences. Written informed consent to participate in this study was provided by the participants' legal guardian/next of kin.

Author contributions

Conflict of interest

After the closure of the research project where the experiments were conducted, A-RS and RK founded a company called Hokema Oy, that attempts to commercialize the game and speech technology developed in this and the preceding research projects. In addition to A-RS, RK, SY, and MK have been listed as inventors. This has had no effect on the results of the current study or their interpretation.

The remaining author declares that the research was conducted in the absence of any commercial or financial relationships that could be construed as a potential conflict of interest.

Publisher's note

All claims expressed in this article are solely those of the authors and do not necessarily represent those of their affiliated organizations, or those of the publisher, the editors and the reviewers. Any product that may be evaluated in this article, or claim that may be made by its manufacturer, is not guaranteed or endorsed by the publisher.

References

Acquah, E. O., and Katz, H. T. (2020). Digital game-based L2 learning outcomes for primary through high-school students: A systematic literature review. *Comput. Educ.* 143:103667. doi: 10.1016/j.compedu.2019.103667

Adlof, S. M., Baron, L. S., Bell, B. A., and Scoggins, J. (2021). Spoken word learning in children with developmental language disorder or dyslexia. *J. Speech Lang. Hear. Res.* 64, 2734–2749. doi: 10.1044/2021_JSLHR-20-00217

Aghlara, L., and Tamjid, N. H. (2011). The effect of digital games on Iranian children's vocabulary retention in foreign language acquisition. *Proc. Soc. Behav. Sci.* 29, 552–560. doi: 10.1016/j.sbspro.2011.11.275

Ahonen, T., Tuovinen, S., and Leppäsaari, T. (2003). *Nopean sarjallisen nimeämisen testi.* Jyväskylä: Niilo Mäki Instituutti & Haukkarannan koulu.

Alt, M., Hogan, T., Green, S., Gray, S., Cabbage, K., and Cowan, N. (2017). Word learning deficits in children with dyslexia. *J. Speech Lang. Hear. Res.* 60, 1012–1028. doi: 10.1044/2016_JSLHR-L-16-0036

Amenedo, E., and Escera, C. (2000). The accuracy of sound duration representation in the human brain determines the accuracy of behavioural perception. *Eur. J. Neurosci.* 12, 2570–2574. doi: 10.1046/j.1460-9568.2000.00114.x

Andreola, C., Mascheretti, S., Belotti, R., Ogliari, A., Marino, C., Battaglia, M., et al. (2021). The heritability of reading and reading-related neurocognitive components: A multi-level meta-analysis. *Neurosci. Biobehav. Rev.* 121, 175–200. doi: 10.1016/j.neubiorev.2020.11.016

Benjamini, Y., and Hochberg, Y. (1995). Controlling the false discovery rate: A practical and powerful approach to multiple testing. *J. R. Stat. Soc. B* 57, 289–300.

Bexkens, A., van den Wildenberg, W. P. M., and Tijms, J. (2015). Rapid automatized naming in children with dyslexia: Is inhibitory control involved? *Dyslexia* 21, 212–234. doi: 10.1002/dys.1487

Bishop, D. V. M., and Snowling, M. J. (2004). Developmental dyslexia and specific language impairment: Same or different? *Psychol. Bull.* 130, 858–886. doi: 10.1037/0033-2909.130.6.858

Boersma, P., and Weenink, D. (2010). *Praat: Doing phonetics by computer [Computer program].* (Version 5.1.45). Available online at: http://www.praat.org/ (accessed January 10, 2017).

Boets, B., Op de Beeck, H. P., Vandermosten, M., Scott, S. K., Gillebert, C. R., Mantini, D., et al. (2013). Intact but less accessible phonetic representations in adults with dyslexia. *Science* 342, 1251–1254. doi: 10.1126/science.1244333

Bruder, J., Leppänen, P. H., Bartling, J., Csépe, V., Démonet, J. F., and Schulte-Körne, G. (2011). Children with dyslexia reveal abnormal native language representations: Evidence from a study of mismatch negativity. *Psychophysiology* 48, 1107–1118. doi: 10.1111/j.1469-8986.2011.01179.x

Carvalhais, L., Limpo, T., Richardson, U., and Castro, S. L. (2020). Effects of the Portuguese graphogame on reading, spelling, and phonological awareness in second graders struggling to read. *J. Writ. Res.* 12, 9–34. doi: 10.17239/JOWR-2020.12.01.02

Downey, D. M., Snyder, L. E., and Hill, B. (2000). College students with dyslexia: Persistent linguistic deficits and foreign language learning. *Dyslexia* 6, 101–111. doi: 10.1002/(SICI)1099-0909(200004/06)6:2<101::AID-DYS154>3.0.CO;2-8

Eden, S., and Shmila, L. (2022). Improving vocabulary in English as a foreign language among students with dyslexia using hybrid technology. *J. Educ. Comput. Res.* 1–21. doi: 10.1177/07356331221117084

Facoetti, A., Luisa Lorusso, M., Paganoni, P., Umiltà, C., and Gastone Mascetti, G. (2003). The role of visuospatial attention in developmental dyslexia: Evidence from a rehabilitation study. *Cogn. Brain Res.* 15, 154–164. doi: 10.1016/S0926-6410(02)00148-9

Feng, G., Yi, H. G., and Chandrasekaran, B. (2019). The role of the human auditory corticostriatal network in speech learning. *Cereb. Cortex* 29, 4077–4089. doi: 10.1093/cercor/bhy289

Fitzgerald, K., and Todd, J. (2020). Making sense of mismatch negativity. *Front. Psychiatry* 11:468. doi: 10.3389/fpsyt.2020.00468

Franciosi, S. J. (2017). The effect of computer game-based learning on fl vocabulary transferability. *J. Educ. Technol. Soc.* 20, 123–133. doi: 10.2307/jeductechsoci.20.1.123

Gu, C., and Bi, H.-Y. (2020). Auditory processing deficit in individuals with dyslexia: A meta-analysis of mismatch negativity. *Neurosci. Biobehav. Rev.* 116, 396–405. doi: 10.1016/j.neubiorev.2020.06.032

Hallam, S., Price, J., and Katsarou, G. (2002). The effects of background music on primary school pupils' task performance. *Educ. Stud.* 28, 111–122. doi: 10.1080/03055690220124551

Häyrinen, T., Serenius-Sirve, S., Korkman, M., Ekebom, U. M., Kivelä, A., Lassander, P., et al. (1999). *Lukilasse. Lukemisen, kirjoittamisen ja laskemisen seulontatestistö peruskoulun ala-asteen luokille 1-6.* Helsinki: Psykologien kustannus.

Helland, T., and Kaasa, R. (2005). Dyslexia in English as a second language. *Dyslexia* 11, 41–60. doi: 10.1002/dys.286

Jarsve, C. F., and Tsagari, D. (2022). Dyslexia and English as a foreign language in norwegian primary education: A mixed methods intervention study. *Cent. Educ. Policy Stud. J.* 12, 155–180. doi: 10.26529/cepsj.1459

Jost, L. B., Eberhard-Moscicka, A. K., Pleisch, G., Heusser, V., Brandeis, D., Zevin, J. D., et al. (2015). Native and non-native speech sound processing and the neural mismatch responses: A longitudinal study on classroom-based foreign language learning. *Neuropsychologia* 72, 94–104. doi: 10.1016/j.neuropsychologia.2015.04.029

Junttila, K., Smolander, A.-R., Karhila, R., Giannakopoulou, A., Uther, M., Kurimo, M., et al. (2022). Gaming enhances learning-induced plastic changes in the brain. *Brain Lang.* 230:105124. doi: 10.1016/j.bandl.2022.105124

Kalyakin, I., González, N., Joutsensalo, J., Huttunen, T., Kaartinen, J., and Lyytinen, H. (2007). Optimal digital filtering versus difference waves on the mismatch negativity in an uninterrupted sound paradigm. *Dev. Neuropsychol.* 31, 429–452. doi: 10.1080/87565640701229607

Karhila, R., Smolander, A. R., Ylinen, S., and Kurimo, M. (2019). "Transparent pronunciation scoring using articulatorily weighted phoneme edit distance," in *Proceedings of the Annual Conference of the International Speech Communication Association, INTERSPEECH, 2019-Septe* (Austria: ISCA), 1866–1870. doi: 10.21437/Interspeech.2019-1785

Karhila, R., Ylinen, S., Enarvi, S., Palomäki, K., Nikulin, A., Rantula, O., et al. (2017). "SIAK - A game for foreign language pronunciation learning," in *Proceedings of the Annual Conference of the International Speech Communication Association, INTERSPEECH, 2017-Augus* (Austria: ISCA), 3429–3430. doi: 10.21437/Interspeech.2017-2046

Krishnan, S., Watkins, K. E., and Bishop, D. V. M. (2016). Neurobiological basis of language learning difficulties. *Trends Cogn. Sci.* 20, 701–714. doi: 10.1016/j.tics.2016.06.012

Leotti, L. A., and Delgado, M. R. (2011). The inherent reward of choice. *Psychol. Sci.* 22, 1310–1318. doi: 10.1177/0956797611417005

Lindeman, J. (1998). *Ala-asteen lukutesti: Käyttäjän käsikirja.* Turku: Turun yliopisto, oppimistutkimuksen keskus.

Liu, T. Y., and Chu, Y. L. (2010). Using ubiquitous games in an English listening and speaking course: Impact on learning outcomes and motivation. *Comput. Educ.* 55, 630–643. doi: 10.1016/j.compedu.2010.02.023

Marco-Pallarés, J., Müller, S. V., and Münte, T. F. (2007). Learning by doing: An fMRI study of feedback-related brain activations. *Neuroreport* 18, 1423–1426. doi: 10.1097/WNR.0b013e3282e9a58c

Murayama, K., Matsumoto, M., Izuma, K., Sugiura, A., Ryan, R. M., Deci, E. L., et al. (2015). How self-determined choice facilitates performance: A key role of the ventromedial prefrontal cortex. *Cereb. Cortex* 25, 1241–1251. doi: 10.1093/cercor/bht317

Näätänen, R., Gaillard, A. W. K., and Mäntysalo, S. (1978). Early selective-attention effect on evoked potential reinterpreted. *Acta Psychol.* 42, 313–329. doi: 10.1016/0001-6918(78)90006-9

Näätänen, R., Lehtokoski, A., Lennes, M., Cheour, M., Huotilainen, M., Iivonen, A., et al. (1997). Language-specific phoneme representations revealed by electric and magnetic brain responses. *Nature* 385, 432–434. doi: 10.1038/385432a0

Nahum, M., and Bavelier, D. (2020). Video games as rich environments to foster brain plasticity. *Handb. Clin. Neurol.* 168, 117–136. doi: 10.1016/B978-0-444-63934-9.00010-X

Nenonen, S., Shestakova, A., Huotilainen, M., and Näätänen, R. (2005). Speech-sound duration processing in a second language is specific to phonetic categories. *Brain Lang.* 92, 26–32. doi: 10.1016/j.bandl.2004.05.005

Oei, A. C., and Patterson, M. D. (2013). Enhancing cognition with video games: A multiple game training study. *PLoS One* 8:e58546. doi: 10.1371/journal.pone.0058546

Park, S., and Lim, J. (2007). Promoting positive emotion in multimedia learning using visual illustrations. *J. Educ. Multimedia Hypermedia* 16, 141–162.

Peters, L., Op de Beeck, H., and De Smedt, B. (2020). Cognitive correlates of dyslexia, dyscalculia and comorbid dyslexia/dyscalculia: Effects of numerical magnitude processing and phonological processing. *Res. Dev. Disabil.* 107:103806. doi: 10.1016/j.ridd.2020.103806

Pettigrew, C. M., Murdoch, B. E., Ponton, C. W., Kei, J., Chenery, H. J., and Alku, P. (2004). Subtitled videos and mismatch negativity (MMN) investigations of spoken word processing. *J. Am. Acad. Audiol.* 15, 469–485. doi: 10.3766/jaaa.15.7.2

Picton, T. W., Alain, C., Otten, L., Ritter, W., and Achim, A. (2000). Mismatch negativity: Different water in the same river. *Audiol. Neuro Otol.* 5, 111–139. doi: 10.1159/000013875

Pulvermüller, F., and Shtyrov, Y. (2006). Language outside the focus of attention: The mismatch negativity as a tool for studying higher cognitive processes. *Prog. Neurobiol.* 79, 49–71. doi: 10.1016/j.pneurobio.2006.04.004

Pulvermüller, F., Kujala, T., Shtyrov, Y., Simola, J., Tiitinen, H., Alku, P., et al. (2001). Memory traces for words as revealed by the mismatch negativity. *Neuroimage* 14, 607–616. doi: 10.1006/nimg.2001.0864

Ramus, F., Marshall, C. R., Rosen, S., and van der Lely, H. K. J. (2013). Phonological deficits in specific language impairment and developmental dyslexia: Towards a multidimensional model. *Brain* 136, 630–645. doi: 10.1093/brain/aws356

Ronimus, M., Eklund, K., Pesu, L., and Lyytinen, H. (2019). Supporting struggling readers with digital game-based learning. *Educ. Technol. Res. Dev.* 67, 639–663. doi: 10.1007/s11423-019-09658-3

Sams, M., Paavilainen, P., Alho, K., and Näätänen, R. (1985). Auditory frequency discrimination and event-related potentials. *Electroencephalogr. Clin. Neurophysiol.* 62, 437–448. doi: 10.1016/0168-5597(85)90054-1

Schulte-Körne, G., Deimel, W., Bartling, J., and Remschmidt, H. (1998). Auditory processing and dyslexia: Evidence for a specific speech processing deficit. *Neuroreport* 9, 337–340. doi: 10.1097/00001756-199801260-00029

Sinkkonen, J., and Tervaniemi, M. (2000). Towards optimal recording and analysis of the mismatch negativity. *Audiol. Neurootol.* 5, 235–246.

Snowling, M. J., and Hulme, C. (2011). Evidence-based interventions for reading and language difficulties: Creating a virtuous circle. *Br. J. Educ. Psychol.* 81, 1–23. doi: 10.1111/j.2044-8279.2010.02014.x

Soroli, E., Szenkovits, G., and Ramus, F. (2010). Exploring dyslexics' phonological deficit III: Foreign speech perception and production. *Dyslexia* 16, 318–340. doi: 10.1002/dys.415

Sussman, E. S. (2007). A new view on the MMN and attention debate. *J. Psychophysiol.* 21:164. doi: 10.1027/0269-8803.21.34.164

Szenkovits, G., and Ramus, F. (2005). Exploring dyslexics' phonological deficit I: Lexical vs sub-lexical and input vs output processes. *Dyslexia* 11, 253–268. doi: 10.1002/dys.308

Tamminen, H., Peltola, M. S., Kujala, T., and Näätänen, R. (2015). Phonetic training and non-native speech perception—New memory traces evolve in just three days as indexed by the mismatch negativity (MMN) and behavioural measures. *Int. J. Psychophysiol.* 97, 23–29. doi: 10.1016/j.ijpsycho.2015.04.020

Temple, E., Deutsch, G. K., Poldrack, R. A., Miller, S. L., Tallal, P., Merzenich, M. M., et al. (2003). Neural deficits in children with dyslexia ameliorated by behavioral remediation: Evidence from functional MRI. *Proc. Natl. Acad. Sci.* 100, 2860–2865. doi: 10.1073/pnas.0030098100

Tiitinen, H., May, P., Reinikainen, K., and Näätänen, R. (1994). Attentive novelty detection in humans is governed by pre-attentive sensory memory. *Nature* 372, 90–92. doi: 10.1038/372090a0

Tremblay, K., Kraus, N., and McGee, T. (1998). The time course of auditory perceptual learning: Neurophysiological changes during speech-sound training. *Neuroreport* 9, 3557–3560. doi: 10.1097/00001756-199811160-00003

Tsai, Y. L., and Tsai, C. C. (2018). Digital game-based second-language vocabulary learning and conditions of research designs: A meta-analysis study. *Comput. Educ.* 125, 345–357. doi: 10.1016/j.compedu.2018.06.020

Turker, S., Seither-Preisler, A., and Reiterer, S. M. (2021). Examining individual differences in language learning: A neurocognitive model of language aptitude. *Neurobiol. Lang.* 2, 389–415. doi: 10.1162/nol_a_00042

van Zuijen, T. L., Plakas, A., Maassen, B. A., Maurits, N. M., and van der Leij, A. (2013). Infant ERPs separate children at risk of dyslexia who become good readers from those who become poor readers. *Dev. Sci.* 16, 554–563. doi: 10.1111/desc.12049

Vandermosten, M., Boets, B., Wouters, J., and Ghesquière, P. (2012). A qualitative and quantitative review of diffusion tensor imaging studies in reading and dyslexia. *Neurosci. Biobehav. Rev.* 36, 1532–1552. doi: 10.1016/j.neubiorev.2012.04.002

Vellutino, F. R., Fletcher, J. M., Snowling, M. J., and Scanlon, D. M. (2004). Specific reading disability (dyslexia): What have we learned in the past four decades? *J. Child Psychol. Psychiatry* 45, 2–40. doi: 10.1046/j.0021-9630.2003.00305.x

Wagner, R. K., and Torgesen, J. K. (1987). The nature of phonological processing and its causal role in the acquisition of reading skills. *Psychol. Bull.* 101, 192–212. doi: 10.1037/0033-2909.101.2.192

Wagner, R. K., Zirps, F. A., Edwards, A. A., Wood, S. G., Joyner, R. E., Becker, B. J., et al. (2020). The prevalence of dyslexia: A new approach to its estimation. *J. Learn. Disabil.* 53, 354–365. doi: 10.1177/0022219420920377

Wang, Y., Mauer, M. V., Raney, T., Peysakhovich, B., Becker, B. L., Sliva, D. D., et al. (2017). Development of tract-specific white matter pathways during early reading development in at-risk children and typical controls. *Cereb. Cortex* 27, 2469–2485. doi: 10.1093/cercor/bhw095

Wechsler, D. (2010). *WISC-IV – Wechsler Intelligence Scale for Children – IV*. Helsinki: Psykologien Kustannus Oy.

Widmann, A., Schröger, E., Tervaniemi, M., Pakarinen, S., and Kujala, T. (2012). Mapping symbols to sounds: Electrophysiological correlates of the impaired reading process in dyslexia. *Front. Psychol.* 3:60. doi: 10.3389/fpsyg.2012.00060

Winkler, I. (2007). Interpreting the mismatch negativity. *J. Psychophysiol.* 21, 147–163. doi: 10.1027/0269-8803.21.34.147

Wittmann, B. C., Bunzeck, N., Dolan, R. J., and Düzel, E. (2007). Anticipation of novelty recruits reward system and hippocampus while promoting recollection. *Neuroimage* 38, 194–202. doi: 10.1016/j.neuroimage.2007.06.038

Wong, P. C. M., Morgan-Short, K., Ettlinger, M., and Zheng, J. (2012). Linking neurogenetics and individual differences in language learning: The dopamine hypothesis. *Cortex* 48, 1091–1102. doi: 10.1016/j.cortex.2012.03.017

Yi, H.-G., Maddox, W. T., Mumford, J. A., and Chandrasekaran, B. (2016). The role of corticostriatal systems in speech category learning. *Cereb. Cortex* 26, 1409–1420. doi: 10.1093/cercor/bhu236

Ylinen, S., and Kurimo, M. (2017). "Kielenoppiminen vauhtiin puheteknologian avulla," in *Oppimisen tulevaisuus*, eds H. Savolainen, R. Vilkko, and L. Vähäkylä (Helsinki: Gaudeamus), 57–69.

Ylinen, S., Junttila, K., Laasonen, M., Iverson, P., Ahonen, L., and Kujala, T. (2019). Diminished brain responses to second-language words are linked with native-language literacy skills in dyslexia. *Neuropsychologia* 122, 105–115. doi: 10.1016/j.neuropsychologia.2018.11.005

Ylinen, S., Smolander, A.-R., Karhila, R., Kakouros, S., Lipsanen, J., Huotilainen, M., et al. (2021). The effects of a digital articulatory game on the ability to perceive speech-sound contrasts in another language. *Front. Educ.* 6:612457. doi: 10.3389/feduc.2021.612457

Ylinen, S., Uther, M., Latvala, A., Vepsäläinen, S., Iverson, P., Akahane-Yamada, R., et al. (2010). Training the brain to weight speech cues differently: A study of Finnish second-language users of English. *J. Cogn. Neurosci.* 22, 1319–1332.