HHS Public Access

Author manuscript

Neuroimage. Author manuscript; available in PMC 2021 August 13.

Published in final edited form as:

Neuroimage. 2021 August 01; 236: 118083. doi:10.1016/j.neuroimage.2021.118083.

Reciprocal relations between reading skill and the neural basis of phonological awareness in 7- to 9-year-old children

Jin Wang[a,*], **Julia Pines**[b], **Marc Joanisse**[c], **James R. Booth**[a]

[a]Department of Psychology and Human Development, Vanderbilt University, Nashville, TN, USA

[b]Neuroscience Program, College of Arts and Sciences, Vanderbilt University, Nashville, TN, USA

[c]Department of Psychology & Brain and Mind Institute, The University of Western Ontario, London, Ontario, Canada

Abstract

By using a longitudinal design and functional magnetic resonance imaging (fMRI), our previous study (Wang et al., 2020) found a scaffolding effect of early phonological processing in the superior temporal gyrus (STG) in 6-year-old children on later behavioral reading skill in 7.5-year-old children. Other than this previous study, nothing is known about longitudinal change in the bidirectional relation between reading skill and phonological processing in the brain. To fill this gap, in the current study, we used the same experimental paradigm as in Wang et al. (2020) to measure children's reading skill and brain activity during an auditory phonological awareness task, but with children who were 7.5 years old at Time 1 (T1) and about 1.5 years later when they were 9 years old at Time 2 (T2). The phonological awareness task included both small grain (i.e., onset) and large grain (i.e., rhyme) conditions. In a univariate analysis, we found that better reading skill at T1 predicted lower brain activation in IFG at T2 for onset processing after controlling for brain activation and non-verbal IQ at T1. This suggests that early reading ability reduces the effort of phonemic access, thus supporting the refinement hypothesis. When using general psychophysiological interaction (gPPI), we found that higher functional connectivity from IFG to STG for rhyme processing at T1 predicted better reading skill at T2 after controlling for reading skill and non-verbal IQ at T1. This suggests that the early effectiveness of accessing rhyme representations scaffolds reading acquisition. As both results did not survive multiple comparison corrections, replication of these findings is needed. However, both findings are consistent with prior studies demonstrating that phonological access in the frontal lobe becomes important in older elementary school readers. Moreover, the refinement effect for onsets is consistent with the hypothesis that learning to read allows for better access of small grain phonology, and the

*Corresponding author jin.wang@vanderbilt.edu (J. Wang).

Author contribution

Jin Wang: Hypothesis formation, Analysis, Manuscript writing;
Julia Pines: Analysis, Manuscript writing;
Marc F. Joanisse: Revision;
James Booth: Experimental design, Hypothesis formation, Revision.

Declaration of Competing Interest

The authors declare that they have no known competing financial interests or personal relationships that could have appeared to influence the work reported in this paper.

scaffolding effect for rhymes supports the idea that reading progresses to larger grain orthography-to-phonology mapping in older skilled readers. The current study, along with our previous study on younger children, indicates that the development of reading skill is associated with (1) the early importance of the quality of the phonological representations to later access of these representations, and (2) early importance of small grain sizes to later development of large grain ones.

Keywords

Phonological awareness; Reading; Longitudinal; fMRI

1. Introduction

Phonological awareness is an individual's ability to represent and access the sound structure of spoken words (Treiman and Zukowski, 1991). This ability is hypothesized to facilitate later reading acquisition because the awareness that a spoken word is composed of small sound units could facilitate a connection between distinct phonemes and discrete letters. We refer to this as the scaffolding hypothesis. In contrast, other researchers suggest that learning to read refines phonological awareness by mapping letters to the acoustically inseparable phonemes (Ziegler and Goswami, 2005). We refer to this as the refinement hypothesis. Examining these two hypotheses is essential in understanding whether individual differences in phonological awareness is the cause or consequence, or both the cause and consequence of individual differences in reading ability in developing children.

Many behavioral studies support the scaffolding hypothesis by showing that phonological awareness training in preschoolers significantly improves their reading skills in first or second grade (e.g., Lundberg et al., 1988). Longitudinal studies have also shown that phonological awareness in kindergarteners predicts their later reading skills in the first few years of school (e.g., Perfetti et al., 1987; Wagner et al., 1997; Hogan et al., 2005; Boets et al., 2010). Some studies (Torgesen et al., 1997; Wagner et al, 1997) have shown that phonological awareness still predicts later reading skill in older elementary children from 2nd to 4th grade and from 3rd to 5th grade. However, the scaffolding effect is smaller in older compared to younger elementary children. Similarly, Boets et al. (2010) and Hogan et al. (2005) showed that phonological awareness in 1st or 2nd grade no longer predicted reading skill in 3rd or 4th grade. In summary, both training and longitudinal studies provide evidence for the scaffolding hypothesis, but this effect seems to decrease or disappear in older elementary years.

In terms of the refinement hypothesis, studies (e.g., Burgess and Lonigan, 1998; Lerner and Lonigan, 2016) have shown that early letter knowledge in 4-year-old children predicts their phonological awareness 6 months or 1 year later. Boets et al. (2010) also showed that early letter knowledge in kindergarten predicted children's later phonological awareness in the first grade. Perfetti et al. (1987) measured children's reading skill and phonological awareness four times while the children were in first grade. They found that earlier word reading skill was predictive of later phonological awareness. Consistent with this finding,

Boets et al. (2010) found that children's reading skill at the end of first grade was predictive of later phonological awareness in third grade. Hogan et al. (2005) found that reading skill in 2nd grade was predictive of phonological awareness in 4th grade, supporting the refinement hypothesis. However, Wagner et al. (1997) studied kindergarteners and followed them for 5 years. They found that early word reading skills did not predict later phonological awareness over an interval of either 1 year or 2 years. They suggested that it was because phonological awareness became stable as children grow older so that word reading has little impact on it. Therefore, behavioral studies suggest there may be a refinement effect of reading on phonological awareness, but inconsistent evidence exists.

Although neuroimaging studies have not addressed scaffolding or refinement, many have explored the relation between reading skill and phonology in the brain. The benefit of using neuroimaging studies is that they could tease apart phonological awareness components by examining brain regions thought to be involved in representing versus accessing phonology. The superior temporal gyrus (STG) is a region often associated with phonological representations involving acoustic and perceptual features (e.g., Leonard and Chang, 2014). The dorsal IFG, however, appears to be associated with accessing and operating on the phonological representations in the STG (Boets et al., 2013; Hagoort, 2014). This distinct function of STG versus IFG for auditory phonological processing is also supported by Myers et al. (2009), which suggests that the STG represents both sensory and perceptual features of phonology, whereas the IFG plays a higher-order role in accessing and computing these representations. Investigating whether phonological representations in STG or access to them in IFG is critical for individual differences in reading skill is a long-standing question (Peterson and Pennington, 2015).

Several neuroimaging studies have explored the relation between reading skill and brain activation in STG and IFG during phonological processing. In 5–6-year-old children, STG was found under-activated for children who were at risk of dyslexia as compared to typically developing children (Raschle et al., 2012). Longitudinal studies (Maurer et al., 2009) have also shown that brain activation in STG during auditory phonological tasks in kindergarteners predicts later reading skills. So, phonological representations in STG appear to be important for reading in young children. In older children and adults, however, the frontal lobe seems to play a more important role in reading skill. Kovelman et al. (Kovelman et al., 2012) observed that children with dyslexia aged 7–13 years old did not activate their frontal lobe during phonological judgments to spoken words whereas all typical readers did. In 10–13-year-old children, Corina et al. (2001) also found that children with dyslexia showed less activation in the frontal cortex during auditory spoken language tasks as compared to typically developing children. Boets et al. (2013) found similar results in dyslexic adults, with subjects having intact phonological representation in the STG but difficulty accessing those representations through the dorsal IFG. In summary, phonological representations in STG seem to be crucial for reading skill in younger children, but phonological access and computations in IFG appears to be more important for reading skills in older children. However, not all research points to this conclusion. Some studies suggest that phonological representations in STG still play an important role in reading skills in older elementary school children. Vandermosten et al. (2019) used multi-voxel pattern analysis and found atypical phonemic representations in STG in 8-year-old children with

dyslexia. Brennan et al. (2013) found a correlation between reading skill and activation in the left STG in 8–12-year-old children. Other research indicates phonological processing in the brain may not be related to reading skills in older children. In 8–13-year-old children, Debska et al. (2019) showed no association of reading skill with activation in STG or IFG during auditory phonological awareness tasks.

The accurate representation and effective access to phonology may depend on grain size, i.e., smaller units at the phoneme level versus larger units at the rhyme level. Cross-sectional studies show that phonemic awareness is more strongly correlated with reading than rhyme awareness (see meta-analysis Melby-Lervåg et al., 2012). Longitudinal studies also found that phonemic awareness in kindergarteners is more powerful in predicting reading gains in the first few years of schooling than rhyme awareness, suggesting that small grain phonological awareness plays a more important role in reading acquisition (e.g., Muter et al., 1998; Hulme et al., 2002; Muter et al., 2004; Castles and Coltheart, 2004). In the only longitudinal study to examine the bidirectional relation of reading skill and phonological processing in the brain, we (Wang et al., 2020) found that the activation in STG for phoneme as well as rhyme judgments in 6-year-old children were predictive of reading skills in 7.5-year-old children. Overall, the literature seems to point to the importance of phonemic awareness in scaffolding reading gains, but the role of large grain sizes seems to be weaker.

As compared to our previous study (Wang et al., 2020) using 6- to 7.5-year-old children, we aimed to investigate the bidirectional relationship between reading skill and phonological processing in the brain in a relatively older cohort aged 7.5–9 years old. Using the same cross-lagged panel design, we examined phonological activation associated with both small (i.e., onset) and large grain (i.e., rhyme) size. We analyzed the brain activity in STG to measure phonological representations and analyzed both brain activity in and functional connectivity with the IFG to measure phonological access to those representations.

To examine the scaffolding hypothesis, we analyzed if brain activity or connectivity at Time 1 (7.5 years old) predicted reading skill at Time 2 (9 years old) after controlling reading skill and other covariates of no interest at Time 1. Based on previous behavioral studies (Torgesen et al., 1997; Wagner et al., 1997) showing that the scaffolding effect lasts until later elementary years, and one previous neural study (Maurer et al., 2009) showing that phonological processing in the brain in kindergarteners predicts later reading skills even in 5th grade, we expected to observe a scaffolding effect in 7.5- to 9-year-old children. However, it is possible that we will not observe a scaffolding effect because studies have shown that this effect decreases or disappears with age (Torgesen et al., 1997; Wagner et al, 1997; Hogan et al., 2005). If we do observe a scaffolding effect in the older children in our study, we expected that brain activation and/or functional connectivity in IFG would play a more important role than STG in scaffolding later reading skill because reading skill seems to be more strongly related to phonological representations in STG in younger children but phonological access in IFG in older children (e.g., Raschle et al., 2012; Dębska et al., 2016; Kovelman et al., 2012; Corina et al., 2001; Boets et al., 2013). As for different grain sizes of phonological processing, we expected that onset processing would play a more important role in predicting later reading skills, because previous behavioral studies have shown that

phonemic awareness is more powerful in predicting reading gains than rhyme awareness (e.g., Muter et al., 1998; Hulme et al., 2002; Muter et al., 2004; Castles and Coltheart, 2004).

To examine the refinement hypothesis, we tested if reading skill at Time 1 (T1, 7.5 years old) predicted brain activity or connectivity during our phonological awareness task at Time 2 (T2, 9 years old) after controlling for brain activity or connectivity and other covariates of no interest at Time 1. Based on the inconsistent findings of previous behavioral studies on older elementary school children (Wagner et al, 1997; Hogan et al., 2005; Boets et al., 2010), it is unclear whether we will observe a refinement effect in 7.5- to 9-year-old children using brain measures in the current study. Brain measures could provide a complementary measure for capturing individual differences in phonological processing. Previous studies have indicated that brain measures can either be a better predictor of reading skills compared to behavioral measures (e.g., Maurer et al., 2009; Wang et al., 2020) or increase the predictive power when combined with behavioral measures (e.g., Kraft et al., 2016; Kuhl et al., 2020). The finding of a refinement effect in the brain of older children would be consistent with theoretical models arguing that reading facilitates the discovery of phonemes (Ziegler and Goswami, 2005), so early reading skill should predict onset processing better than rhyme processing in the brain. If we find that reading refines phonemic processing, this should be larger for accessing phonology in IFG compared to representing phonology in STG, consistent with previous neuroimaging studies showing reading skill is more strongly correlated with IFG in older children compared to STG in younger children (e.g., Raschle et al., 2012; (Dębska et al., 2016)(Kovelman et al., 2012) Corina et al., 2001; Boets et al., 2013).

In summary, evidence suggests that we may observe a scaffolding or refinement effect in the current study on 7.5- to 9-year-old children. If we do observe these effects, they should be stronger for detecting onsets that require phonemic processing and for the IFG which is involved in accessing posterior phonological representations.

2. Method

2.1. Participants

Fifty-nine monolingual English-speaking children (32 females, mean age = 7.3, range 7.0–8.2 years old at Time 1, mean age = 9.2, range 9.0–9.9 years old at Time 2) were included in this study. All the participants were recruited in the Austin metropolitan area. The Institutional Review Board at The University of Texas at Austin approved all of the experimental procedures.

Parents of our participants were asked to complete an exclusionary survey and a developmental questionnaire. Then, participants completed several screening tests that included 5-handedness questions in which the children had to pretend to write, erase, pick, open, and throw something, as well as the Diagnostic Evaluation of Language Variation (DELV) Part 1 Language Variation Status (Seymour et al., 2003). Participants were included if they met the following criteria: (1) right handed, defined as completing at least 3 out of the 5 tasks in the 5-handedness questions with their right hand when they entered the project; (2) a mainstream American English speaker, defined as having at or above the

following criteria: 9 out of 15 for 7-year-olds, 11 out of 15 for 8-year-olds, 12 out of 15 for 9 and 10-year-olds mainstream English responses on the DELV Part I Language Variation Status test; (3) no learning, neurological or psychiatric disorders, including Attention Deficit Hyperactivity Disorder (ADHD), according to the developmental history questionnaire completed by the parents; and (4) normal hearing and normal or corrected-to-normal vision as reported by their parents.

These children also completed a series of standardized tests to assess their language ability, non-verbal IQ, phonological awareness, and reading skill. Language ability was measured by the Core Language Scale on the Clinical Evaluation of Language Fundamentals, Fifth Edition (CELF- 5, Wiig et al., 2013). Non-verbal IQ was measured by the Kaufman Brief Intelligence Test, Second Edition (KBIT-2, Kaufman and Kaufman, 2004). All children had normal IQ and language ability as indexed by having a standardized score greater than 70 for both CELF-5 Core Language Scale and KBIT-2 Non-verbal IQ at T1. Non-verbal IQ was also used as a control variable in the main analysis. Phonological awareness and reading, which are the two variables of interest, were also measured using standardized tests at both T1 and T2. Phonological awareness was measured by three subtests on the Comprehensive Test of Phonological Processing (CTOPP-2, Wagner et al., 2013), which included elision, blending sounds and phoneme isolation. The raw score of phonological awareness is the sum of the scaled scores on the three subtests. The seeming decrease in phonological awareness composite scores does not represent a decrease in their skills but indicates that our group of children developed slower than their age-matched cohort. Reading ability was measured by the raw scores of the Woodcock-Johnson III Test of Achievement Letter-Word Identification subtest (Woodcock et al., 2001). Children were required to read the visually presented letters and words out loud. The raw score of Letter-Word Identification is the number of items correctly read by children. Descriptive statistics for the standardized test scores are shown in Table 1.

There were 99 children who originally enrolled in this study for both T1 and T2 sessions with full runs. One was excluded due to left-handedness. Nine were excluded due to not being mainstream English speakers. Ten were excluded after screening for movement (see criteria in the 2.4 data analysis section). Nineteen were excluded because they did not meet the accuracy criteria for performing the fMRI task (see criteria in the 2.2 procedure section). In the end, 59 subjects were included in our final analysis.

2.2. Procedure

The auditory phonological judgment task was an event-related design. Fig. 1 illustrates a description of the task procedure. During each trial, children heard two auditory stimuli presented sequentially and binaurally through earphones. There were four conditions of the pairs of stimuli: onset, rhyme, non-match, and perceptual (frequency modulated noise), examples of which can be seen in Table 2. Participants were asked, “do the two words share the same sound”. They were instructed to respond to all trials as quickly and accurately as possible with the right index finger indicating a yes response in the onset, rhyme and perceptual conditions, and the right middle finger for a no response in the non-match condition. A blue circle remained on the screen during the auditory stimuli presentation

and it turned to yellow 1000 ms before the trial ended to remind participants to respond. The duration of each word was between 500 and 700 milliseconds (ms) followed by a brief period of silence, with the second word beginning 1000 ms after the onset of the first. The duration of the response interval was 1800ms. There were 24 trials for each of the four conditions, divided into two runs. The four conditions were pseudo-randomized so there were no more than 5 of the same responses in a row. To aid in convolving the hemodynamic response, inter-trial intervals were jittered by randomly adding 0, 450 or 900 ms for each trial, in equal proportions for the first run. For the second run, jitters of 0, 375 or 750 ms were similarly added to the trials. Each run lasted about 3 min.

The auditory word conditions were designed according to the following standards (see Table 2 for examples). For the onset condition, the word pairs only shared the same initial phoneme (corresponding to one letter of its written form). For the rhyme condition, the word pairs shared the same vowel and final phoneme/cluster (2–3 letters at the end of its written form). For the non-match condition, there were no shared phonemes (or letters of its written form). All the words were monosyllabic. Every paired word had no semantic association based on the University of South Florida Free Association Norms (Nelson et al., 1998). There were no significant differences between conditions in duration [Onset vs. Rhyme: ps > 0.42; Onset vs. Non-match: ps > 0.58; Rhyme vs. Non-match: ps > 0.54], phonotactic frequency (Vitevitch and Luce, 2004) [Onset vs. Rhyme: ps > 0.49; Onset vs. Non-match: ps > 0.49; Rhyme vs. Non-match: ps > 0.48], word frequency (Balota et al., 2007) [Onset vs. Rhyme: ps > 0.17; Onset vs. Non-match: ps > 0.17; Rhyme vs. Non-match: ps > 0.38], part of speech (Balota et al., 2007), and phonological or orthographic consistency (Bolger et al., 2008) [Onset vs. Rhyme: ps > 0.13; Onset vs. Non-match: ps > 0.05; Rhyme vs. Non-match: ps > 0.20]. Neither irregular spelling forms nor inflected forms of words were used.

In order to make sure the participants understood the task and to acclimate them to the scanner environment, they were required to complete the same task with different stimuli in the mock scanner and a short practice just before the fMRI scanning session.

Participants who scored within an acceptable accuracy range and had no response bias were included in our analysis (see sample size change in the 2.1 participants section). We included children who scored greater than 50% on the perceptual and rhyme conditions suggesting that they were engaged during the task, and who had an accuracy difference between the rhyme and non-match condition lower than 40% suggesting no obvious response bias. The accuracies for each condition during our auditory phonological task inside the scanner at both Time 1 and Time 2 are shown in Table 3.

2.3. Data acquisition

Participants lay in the scanner with a response button box placed in their right hand. To keep participants focused on the task so that they would respond in time, visual stimuli were projected onto a screen, viewed via a mirror attached to the inside of the head coil. Participants wore earphones to hear the auditory stimuli and two ear pads were used to attenuate the scanner noise. The two phonological task runs were counterbalanced across participants.

Images were acquired using 3.0 T Skyra Siemens scanner with a 64-channel head coil. The blood oxygen level dependent (BOLD) signal was measured using a susceptibility weighted single-shot echo planar imaging (EPI) method. Functional images were acquired with multiband EPI (TE=30 ms, flip angle=80, matrix size=128 × 128, FOV=256 mm^2, slice thickness=2 mm without gap, number of slices=56, TR=1250 ms, Multi-band accel. factor=4, voxel size=2 × 2 × 2 mm). A high resolution T1 weighted MPRAGE scan was acquired with the following scan parameters: TR=1900ms, TE=2.34ms, matrix size=256 × 256, field of view=256 mm^2, slice thickness=1 mm, number of slices=192.

2.4. Data analysis

fMRI data was analyzed using Statistical Parametric Mapping (SPM12, http://www.fil.ion.ucl.ac.uk/spm). First, all functional images were realigned to their mean functional image across runs. The anatomical image was then segmented and warped to a pediatric tissue probability map template to get the transformation field. An anatomical brain mask was created by combining the segmentation products (i.e., grey, white, and cerebrospinal fluid), and then applied to its original anatomical image to produce a skull-stripped anatomical image. Then, the mean functional image and all functional images were co-registered to the skull-stripped anatomical image. Then, all the functional images were normalized to a pediatric template by applying the transformation field to them and re-sampled with a voxel size at 2 × 2 × 2 mm. We created this pediatric tissue probability map template using CerebroMatic (Wilke, et al., 2017), a tool that makes SPM12 compatible pediatric templates with user-defined age, gender, and magnetic field. We inputted the following information into CerebroMatic: the unified segmentation parameters described in Wilke et al. (2017), which were estimated from 1919 participants (https://www.medizin.uni-tuebingen.de/kinder/en/research/neuroimaging/software/) and user defined age as 7–10.5 years old with one-month intervals, gender as two females and two males at each age interval and magnetic field strength as 3T, resulting in a sample of 172 for our pediatric template. After normalization, smoothing was applied to all the functional images with 6 mm isotropic Gaussian kernel.

To reduce movement effects on brain signal, Art-Repair (Mazaika et al., 2009, http://cibsr.stanford.edu/tools/human-brain-project/artrepair-software.html) was used to identify outlier volumes, defined as those with volume-to-volume head movement exceeding 1.5 mm in any direction, head movement greater than 5 mm in any direction from the mean functional image across runs, or deviations of more than 4% from the mean global signal intensity. The outlier volumes were repaired by interpolation by the nearest non-outlier volumes. Subjects included in our study had no more than 10% of the volumes repaired in each run and no more than 6 consecutive volumes repaired in each run. Six motion parameters estimated in the realignment step were entered in the first level modeling as regressors and the repaired volumes were deweighted.

Statistical analyses at the first level were calculated using an event-related design with the four conditions (i.e., onset, rhyme, non-match and perceptual) in each run at each timepoint as conditions of interest. A high pass filter with a cutoff of 128s and an SPM default mask threshold of 0.5 were applied. All experimental trials were included in the analysis. Word

and perceptual pairs were treated as individual events for analysis and modeled using a canonical hemodynamic response function (HRF). Contrast maps were generated for onset > perceptual and rhyme > perceptual at both T 1 and T2 for each participant at the first level analysis. These two contrasts were used to examine how small (i.e., onset) and large (i.e., rhyme) grain sizes of phonological processing played a role in their relations to reading skill. We used one sample *t*-tests at group level analysis to display the activation maps within the whole brain mask for each contrast. We also calculated the T2 > T1 activation maps for onset > perceptual and rhyme > perceptual within each subject at the first level analysis and used one sample *t*-test at the group level to display the brain activation changes over time during onset and rhyme processing (see Fig. 2, Table 4).

Statistical significance for the group level analysis within the whole brain mask (172,512 voxels) was defined using Monte Carlo simulations using AFNI's 3dClustSim program (see http://afni.nimh.nih.gov/). 3dClustSim carries out a 10,000 iteration Monte Carlo simulation of random noise activations at a particular voxel-wise alpha level within a masked brain volume. Following the suggestions made by Eklund et al. (2016) regarding the inflated statistical significance achieved using some packages, we used 3dFWHMx to calculate the smoothness of the data for every single participant, using a spatial autocorrelation function, and then averaged those smoothness values across all participants (ACF =0.48, 4.58, 13.12). This average smoothness value was then entered into 3dClustSim to calculate the cluster size needed for significance. The threshold for the size of a significant cluster within the whole brain mask was 88 voxels at a voxel-wise threshold at $p < 0.001$ uncorrected and cluster-wise threshold at $p < 0.05$ corrected.

Two anatomical masks were used to isolate our regions of interest (ROIs). The posterior left STG was defined as the posterior half of STG with $y < -24$ (Hickok and Poeppel, 2000), while the dorsal left IFG was defined as the opercular part of the left IFG (Boets et al., 2013; Ramus, 2014) by using the anatomical automatic labeling (AAL) atlas template from WFU PickAtlas toolbox(http://www.nitrc.org/projects/wfu_pickatlas).

To examine the scaffolding hypothesis, the top 100 voxels showing maximal activation (regardless of significance) for the contrast of onset>perceptual or rhyme>perceptual at T1 were selected based on their contrast t-maps for every participant within the anatomical mask of the posterior left STG. Beta values were then extracted from these individualized ROIs using Marsbar (http://marsbar.sourceforge.net/tutorial/index.html). After that, a hierarchical regression analysis was run in SPSS, with non-verbal IQ and reading skill at T1 entered into the model as covariates of no interest and brain activation of onset > perceptual at T1 entered as the covariate of interest. The dependent measure was reading skill at T2 (see Table 5). In this way, we examined whether the representational quality of phonemic awareness scaffolds later reading. The same analysis was done using the contrast of rhyme > perceptual at T1 to examine whether the representational quality of rhyme awareness scaffolds later reading. The overlap among participants' individualized ROI within the mask of the posterior left STG for onset > perceptual at T1 and rhyme > perceptual at T1 are plotted in Fig. 3(A) on the left.

To examine the refinement hypothesis, the top 100 voxels showing maximal activation (regardless of significance) for the contrast of onset > perceptual or rhyme > perceptual at T2 were selected based on their contrast t-maps for every participant within the anatomical mask of the posterior left STG. Beta values were extracted from these individualized ROIs using Marsbar (http://marsbar.sourceforge.net/tutorial/index.html). After that, a hierarchical regression analysis was run in SPSS, with non-verbal IQ and brain activation of onset > perceptual at T1 entered into the model as covariates of no interest and reading skill at T1 entered as the covariate of interest. The dependent measure was brain activation of onset > perceptual at T2 (see Table 6). In this way, we examined whether early reading skill refines the later representational quality of phonemic awareness. The same analysis was done using the contrast of rhyme > perceptual to examine whether early reading skill refines later representational quality of rhyme awareness. The overlap among participants' individualized ROI within the mask of posterior STG for onset > perceptual at T2 and rhyme > perceptual at T2 are plotted in Fig. 3(A) on the right.

Parallel univariate analyses, using the mask of the opercular part of left IFG instead of using the posterior left STG, were conducted to examine the scaffolding and refinement hypotheses between reading and phonological access for different grain sizes. The overlap of individualized ROI at both T1 and T2 within the opercular part of IFG are plotted in Fig. 3(B). Because four brain activation models were tested, Bonferroni correction ($p < 0.05/4 = 0.0125$) was applied to determine significance of brain activation results in order to correct for multiple comparisons.

In addition to the univariate analyses, we used a general psychophysiological interaction analysis (gPPI, http://www.nitrc.org/projects/gppi), an approach allowing an investigation of the connectivity strength from one brain area to other areas under a certain experimental condition (McLaren et al., 2012), to evaluate the functional connectivity of IFG with STG during either onset or rhyme processing at T1 or T2.

To assess the functional connectivity between IFG and STG at T1, the top 100 voxels showing maximal activation (regardless of significance) for the contrast of onset>perceptual or rhyme>perceptual at T1 in the opercular part of the left IFG (see Fig. 3B on the left) were used as the seed region. The timeseries from the seed region was extracted. The following regressors were then entered into a general linear model (GLM) in the individual level analysis: the timeseries from the seed region, the 8 experimental parameter regressors, the 8 PPI regressors of the interaction, and the 6 motion regressors of head movement. The 8 experimental parameters were formed by onset times of the onset, rhyme, non-match, and perceptual conditions in run1 at T1 and run2 at T1, respectively. The contrast of onset > perceptual or rhyme > perceptual was defined to produce an individual level functional connectivity map. A one-sample *t*-test group level analysis was performed to show group level functional connectivity from the seed region (i.e., IFG) to the posterior left STG for either onset or rhyme processing at T1 (see Fig. 4A). 3dClustSim was used to determine the significance of a cluster. The threshold for the size of a significant cluster within the posterior left STG mask (1,132 voxels) was 9 voxels using a voxel-wise threshold at $p < 0.001$ uncorrected and cluster-wise threshold at $p < 0.05$ corrected.

To examine the scaffolding hypothesis, we then selected the top 100 voxels within the posterior left STG based on the t-map of the PPI contrast (i.e., onset > perceptual or rhyme > perceptual) at T1 from each participant as the individualized connectivity ROIs. We used Marsbar to extract the PPI betas associated with each condition at T1 from these ROIs. The functional connectivity from IFG to STG for each participant for either onset processing or rhyme processing was calculated by using the PPI betas for onset minus the PPI betas for perceptual or by using the PPI betas for rhyme minus the PPI betas for perceptual at T1. After that, a hierarchical regression analysis was run in SPSS, with non-verbal IQ and reading skill at T1 entered into the model as covariates of no interest and functional connectivity of onset > perceptual at T1 entered as the covariate of interest. The dependent measure was reading skill at T2 (see Table 7). In this way, we examined whether the effectiveness of accessing phonemic representations scaffolds later reading. The same analysis was performed using the contrast of rhyme > perceptual at T1 to examine whether the effectiveness of accessing rhyme representations scaffolds later reading. The overlap among participants' individualized connectivity ROIs within the mask of the posterior left STG for onset > perceptual at T1 and rhyme > perceptual at T1 are plotted in Fig. 4(B).

To evaluate the functional connectivity between IFG and STG at T2, the top 100 voxels showing maximal activation (regardless of significance) for the contrast of onset>perceptual or rhyme>perceptual at T2 in the opercular part of the left IFG (see Fig. 3B on the right) were used as the seed region. The same PPI GLM was performed except that the 8 experimental parameters were formed by onset times of the onset, rhyme, non-match, and perceptual conditions in run1 at T2 and run2 at T2, respectively, using T2 data. A one-sample *t*-test group level analysis was also conducted and Fig. 4C shows the group level functional connectivity from the seed region (i.e., IFG) to the posterior left STG for either onset or rhyme processing at T2. 3dClustSim was used to determine the significance of a cluster. The threshold for the size of a significant cluster within the posterior left STG mask (1,132 voxels) was 9 voxels using a voxel-wise threshold at $p < 0.001$ uncorrected and cluster-wise threshold at $p < 0.05$ corrected.

To examine the refinement hypothesis, we then selected the top 100 voxels within the posterior left STG based on the t-map of the PPI contrast (i.e., onset > perceptual or rhyme > perceptual) at T2 from each participant as the individualized connectivity ROIs. We used Marsbar to extract the functional connectivity from IFG to STG for each participant for either onset processing or rhyme processing at T2. Because we needed to control for the autoregressive effect, we used the same seed region to run the gPPI GLM for each participant but with T1 data and extracted the functional connectivity between the seed and the ROIs at T1. Paired sample *t* tests showed that functional connectivity from IFG to STG increased significantly over time for both onset [mean_T1 = 0.72, mean_T2 = 5.37, $t(58) = -11.048$, $p < .001$] and rhyme processing [mean_T1 = 0.15, mean_T2 = 3.43, $t(58) = -10.138$, $p < .001$]. After that, a hierarchical regression analysis was run in SPSS, with non-verbal IQ and functional connectivity for onset > perceptual at T1 entered into the model as covariates of no interest and reading skill at T1 entered as the covariate of interest. The dependent measure was functional connectivity for onset > perceptual at T2 (see Table 8). In this way, we examined whether early reading skill refines the effectiveness of accessing phonemic representations. The same analysis was performed using the contrast of

rhyme > perceptual to examine whether early reading refines the effectiveness of accessing rhyme representations. The overlap among participants' individualized connectivity ROI within the mask of the posterior left STG for onset > perceptual at T2 and rhyme > perceptual at T2 are plotted in Fig. 4(D). Because two functional connectivity models were tested, Bonferroni correction ($p < 0.05/2 = 0.025$) was applied to determine significance of functional connectivity results in order to correct for multiple comparisons.

In addition to the analysis using top 100 voxels as individualized ROIs, we also used the top 50 and 150 voxels to examine the stability of the results. All results remained the same, with larger effects when using fewer top voxels. In addition to the brain data analyses, behavioral measures were used to examine the scaffolding and the refinement hypotheses. The raw score of phonological awareness from CTOPP-2, the in-scanner task performance for onset and rhyme conditions were used as indices for phonological awareness skills. Then the same hierarchical regression analyses were conducted to examine the scaffolding and the refinement effects, respectively. Because we used 3 indices for phonological awareness to examine each hypothesis, respectively, Bonferroni correction ($p < 0.05/3 = 0.017$) was used to determine the significance of behavioral findings.

3. Results

3.1. Brain activation results

3.1.1. The relation of earlier brain activation in STG to later reading skill—To examine the bidirectional relation of different grain sizes in phonological processing and reading skill, we analyzed both onset and rhyme processing in the brain. The regression analysis showed that brain activation in the posterior left STG for onset > perceptual and rhyme > perceptual did not significantly predict reading skill at T2 after controlling for the reading skill and nonverbal IQ at T1 (see Table 5). Fig. 5 (A) and (B) shows the scatterplots for the relation between brain activation for onset and rhyme processing in STG at T1 and the residuals of reading skill at T2 after controlling for reading skill and non-verbal IQ at T1.

3.1.2. The relation of earlier brain activation in IFG to later reading skill—The regression analysis showed that brain activation in the opercular part of IFG for both onset > perceptual and rhyme > perceptual did not significantly predict reading skill at T2 after the effects of reading skill and nonverbal IQ at T1 were accounted for (see Table 5). Fig. 5 (C) and (D) shows the scatterplots for the relation between brain activation for onset and rhyme processing in IFG at T1 and the residuals of reading skill at T2 after controlling for reading skill and non-verbal IQ at T1.

3.1.3. The relation of earlier reading skill to later brain activation in STG—The regression analysis showed that reading skill did not significantly predict onset > perceptual in the posterior STG at T2 after the effects of brain activation and non-verbal IQ at T1 were accounted for. The regression analysis also showed that reading skill did not significantly predict rhyme > perceptual in the posterior STG at T2 after the effects of brain activation and non-verbal IQ at T1 were accounted for (see Table 6). Fig. 6 (A) and (B) shows the relation between reading skill at T1 and the residuals of brain activation for onset and

rhyme processing in the posterior left STG at T2 after controlling for brain activation and non-verbal IQ at T1.

3.1.4. The relation of earlier reading skill to later brain activation in IFG—

The regression analysis showed that reading skill was related to onset > perceptual in the opercular part of IFG at T2 ($\Delta R^2 = 0.079$, $p = 0.024$) after the effects of brain activation and non-verbal IQ at T1 were accounted for (see Table 6). However, this finding did not survive multiple correction for the 4 brain activation models ($p < 0.05/4=0.0125$). In addition, the regression analysis showed that reading skill did not significantly predict rhyme > perceptual in the opercular part of IFG at T2 after the effects of brain activation and non-verbal IQ at T1 were accounted for (see Table 6).

Fig. 6 (C) and (D) shows the relation between reading skill at T1 and the residuals of brain activation for onset and rhyme processing in the dorsal IFG at T2 after controlling for brain activation and non-verbal IQ at T1.We observed a negative correlation between initial reading skill and activation in the opercular part of IFG at T2 in Fig. 6 (C). Because 3 data points appear to have extreme values, we applied a weighted symmetric Winsorization (Dixon, 1960) to replace the one outlier in the reading measure and the two outliers in the brain activation measure with their nearest neighbors. Then we re-calculated our analysis. We found that our finding was the same (beta = −0.264, $\Delta R^2 = 0.059$, $p = 0.049$). We further plotted the brain activation in the opercular part of IFG at both T1 and T2 for both high and low initial reading groups using a median split (see Fig. 7). This illustrates that brain activation increased over time and that children with higher initial reading skill had a smaller change from T1 to T2 in the activation of the opercular part of IFG for onset processing.

3.2. Brain connectivity results

3.2.1. The relation of earlier functional connectivity between IFG and STG to later reading skill—

In the examination of the scaffolding hypothesis (see Table 7), we found that functional connectivity of IFG with STG for onset processing at T1 did not predict reading skill at T2 after controlling reading skill and non-verbal IQ at T1. However, we did find that functional connectivity of IFG with STG for rhyme processing at T1 predicted reading skill at T2 after controlling reading skill and non-verbal IQ at T1 ($\Delta R^2 = 0.023$, $p = 0.029$). However, this finding did not survive multiple correction for 2 scaffolding brain connectivity models ($p < 0.05/2=0.025$). Fig. 8 (A) and (B) shows the scatterplots for the relation between functional connectivity of IFG with STG for onset or rhyme processing at T1 and the residuals of reading skill at T2 after controlling for reading skill and non-verbal IQ at T1. Because one datapoint appeared to be an outlier, we applied a symmetric Winsorization (Dixon, 1960) to replace the outlier in the functional connectivity measure with its nearest neighbor. Then we re-calculated our analysis. We found that our finding was the same (beta = .146, $\Delta R^2 = 0.020$, $p = 0.038$).

3.2.2. The relation of earlier reading skill to later functional connectivity between IFG and STG—

In the examination of the refinement hypothesis (see Table 8), we found that reading skill at T1 did not predict later functional connectivity of IFG

with STG for onset processing after controlling for functional connectivity and non-verbal IQ at T1. We also did not find that reading skill at T1 predicted later functional connectivity of IFG with STG for rhyme processing after controlling for functional connectivity and non-verbal IQ at T1. Fig. 9 (A) and (B) shows the scatterplots for the relation between reading skill at T1 and the residuals of brain connectivity for onset and rhyme processing of the opercular part of the left IFG with the posterior left STG at T2 after controlling for brain connectivity and non-verbal IQ at T1.

3.3. Behavioral results

3.3.1. The relation of earlier phonological awareness performance to later reading skill—In parallel with the analysis of brain data, the same regression analyses, using task performance inside the scanner and performance on standardized testing, were conducted to examine the scaffolding hypothesis. We found that the accuracy for both the onset and rhyme conditions at T1 did not significantly predict reading skill at T2, after controlling for the reading skill and non-verbal IQ at T1 [onset: $\Delta R^2 = 0.015$, $p = .073$; rhyme: $\Delta R^2 = 0.012$, $p = .119$]. Moreover, parallel regression analysis was calculated using the composite score of Phonological Awareness (PA) on the CTOPP-2. We found that PA at T1 was not predictive of reading skill at T2 after controlling for reading skill and non-verbal IQ at T1 [$\Delta R^2 = 0.002$, $p = .518$], suggesting no scaffolding effect.

3.3.2. The relation of earlier reading skill to later phonological awareness performance—In the examination of the refinement hypothesis, we found that reading skill at T1 did not significantly predict accuracy of the onset or rhyme conditions inside the scanner at T2 after controlling for accuracies at T1 for the onset or rhyme conditions, and nonverbal IQ at T1 [onset: $\Delta R^2 = 0$, $p = .906$; rhyme: $\Delta R^2 = 0.003$, $p = .672$]. In contrast, when using the composite score of Phonological Awareness (PA) on the CTOPP-2, we found that reading skill at T1 predicted PA at T2 after controlling for PA and non-verbal IQ at T1 [$\Delta R^2 = 0.046$, $p = .020$]. However, this result did not survive multiple correction for 3 refinement behavioral models ($p < 0.05/3 = 0.017$). Thus, this result only shows weak evidence towards a refinement effect of T1 reading skill on T2 phonological awareness performance. This finding, however, is consistent with the finding we showed for brain activation analyses.

4. Discussion

The objective of the current study was to investigate the bidirectional relationship between reading skill and phonological processing in the brain in a longitudinal study of children aged 7.5 to 9 years old using functional magnetic resonance imaging (fMRI). When studying younger children from 6 to 7.5 years old, our previous study (Wang et al., 2020) only showed a scaffolding effect of phonological representations in STG on later reading skill and this scaffolding effect occurred for both onset and rhyme. In the current study with older children, however, in support of the refinement hypothesis, we found weak evidence for early reading skill in 7.5-year-old children to predict their brain activation for onset processing in the opercular part of the left IFG one and a half years later. This effect was specific to accessing individual phonemes in frontal cortex, as we did not find

that early reading refined later brain activation in the opercular part of IFG for rhyme processing or brain activation in the posterior STG for either onset or rhyme processing. In the current study, we additionally used general psychophysiological interaction (gPPI) analysis to evaluate IFG's effectiveness in accessing phonological representations stored in the posterior left STG. In contrast to the univariate analysis, we did not find a refinement effect. However, we found weak evidence for scaffolding in which functional connectivity of the opercular part of IFG with the posterior left STG for rhyme processing in 7.5-year-old children predicted their reading skill one and a half years later. This effect was specific to rhyme processing as we did not find that functional connectivity of IFG with STG for onset processing predicted later reading skill. Replication is needed as both findings did not survive multiple correction.

Our current study suggests a reciprocal relation between reading skill and phonological access in older elementary school children. Earlier reading refines later phonological access to phonemes, whereas the effectiveness of accessing rhyme representations scaffolds later reading. Together with the findings from our previous study on younger children (Wang et al., 2020) showing only a scaffolding effect of phonological representations in STG on later reading skill, our current research suggests a developmental progression in the relation between phonological processing and reading skill (see the schematic diagram in Fig. 10). That is, phonological representations in STG scaffold reading acquisition early in development, whereas phonological access in IFG is refined by and scaffolds reading skill later in development. Moreover, as children develop, the access of larger grain sizes of phonology become more critical in reading acquisition. Because both of our major findings in the current study did not survive multiple comparison correction, replication of the findings is needed in future studies. In addition, because the 7.5-year-old children in the current study only overlapped with a portion of the participants in our previous study (Wang et al., 2020), future neuroimaging research using the same children across multiple time points is needed to confirm this developmental change.

The developmental progressions in the bidirectional relation between phonological awareness and reading skill as indicated by our studies are generally consistent with the literature. First, our findings suggest a transition from the importance of phonological representation in the STG to phonological access in the IFG as determinants of reading skill as children develop from the early to middle elementary years. This transition is consistent with previous neural studies showing that reading skill most strongly related to activation during phonological processing in the STG in younger children, but in the IFG in older children (e.g., Raschle et al., 2012; Dębska et al., 2016; Kovelman et al., 2012; Corina et al., 2001; Boets et al., 2013). This transition is also consistent with the argument that young children diagnosed with or at risk of dyslexia usually have impaired phonological representations, whereas older children or adults with dyslexia have impaired phonological access but intact phonological representations (Boets, 2014). Second, our studies suggests that early reading acquisition is marked by scaffolding and that refinement only appears later in development. This is consistent with the longstanding argument that beginning readers rely heavily on phonological awareness to establish the letter-to-sound mappings (Chall, 1983). As decoding becomes more fluent, the scaffolding effect decreases in middle

elementary school. This fluency facilitates the development of sensitivity to the phonemes that the letters represent, which is essential for phonemic awareness.

The third developental progression that our studies suggest is that scaffolding is marked by the increasing reliance on large grain sizes. This change aligns with the theory of reading development by Frith (1985), which argues that reading progresses from the alphabetic stage, relying on small grain letter-to-phoneme mapping, to the orthographic stage, employing larger grain orthographic-to-phonology mapping. This progression is consistent with our previous studies (Wang et al., 2018; Wang et al., 2021) showing that better reading was associated with activation during auditory processing in the ventral occipitotemporal cortex (vOT). However, this relation was shown in the posterior vOT implicated in processing letters in younger children and in the anterior vOT implicated in processing rimes in older children. This engagement from posterior to anterior vOT suggests a transition of the importance of small to large grain sizes for better reading. Although previous longitudinal behavioral studies showed that phonemic awareness is a more powerful predictor of later reading skill than rhyme awareness (e.g., Muter et al., 1998; Hulme et al., 2002; Muter et al., 2004; Castles and Coltheart, 2004), these studies have conducted on early elementary school children. Because rhyme awareness develops earlier than phonemic awareness (Anthony and Francis, 2005), there could be a ceiling effect in the performance of rhyme awareness tasks in older elementary school children. Thus, the developmental transition in the role of phonological grain sizes in reading skill may be hard to detect using behavioral measurements. By using brain measurements in a longitudinal study, our study provides causal evidence for the importance of accessing rhymes in the development of reading skill.

The current study provides the first evidence that reading skill refines phonological access in the brain, which is also supported by our finding that reading skill predicted behavioral gains in phonological awareness. These results are consistent with two previous longitudinal behavioral studies on children at similar age (Hogan et al., 2005; Boets et al., 2010), in which the authors found that reading skill at grade 1 or 2 significantly predicted children's phonological awareness at grade 3 or 4. Both the brain and behavioral predictions did not survive multiple correction, which could be due to a lack of power. The sample size in the current study, although large by neuroimaging standards, is smaller compared to previous behavioral research which included hundreds of children (e.g., Hogan et al., 2005). Although the sample size in Boets et al. (2010) was similar to ours, their measure of phonological awareness was more complex, which may have increased its sensitivity in detecting individual differences. Our finding of a refinement effect in older children is inconsistent with another previous longitudinal study (Wagner et al., 1997), which did not find a refinement effect in children of a similar age, from 2nd to 4th grade. This may be because in the Wagner et al. (1997) study, children's phonological awareness scores were very stable [$r(216) = 0.94$ for 2nd and 4th grade phonological awareness scores], possibly due to repeated tests for 5 years since kindergarten. Thus, there was little room for reading skill to account for variance on phonological awareness after controlling for initial skill. The correlation of phonological awareness scores in the current study between the two time points was more moderate [$r(56) = 0.71$, $p < 0.001$], perhaps allowing us to detect a refinement effect.

Our use of brain measures provides additional insight by showing that reading refined later brain activation during phonemic processing in the opercular part of IFG rather than the posterior left STG. According to the Memory, Unification, and Control (MUC) model by Hagoort (2014) and the study by Myers et al., (2009), regions in the frontal cortex such as the dorsal IFG are crucial for accessing phonological representations, whereas the temporal cortex subserves knowledge representations that have been laid down in memory during acquisition. We did not find a refinement effect in posterior STG, so our neural results suggest that the refinement effect of reading skill is only on the access of phonology, and not on the phonological representations themselves. We may not have found a refinement effect of reading skill on phonological representations in STG in either early or late elementary years because refinement of representations appears to occur earlier in preschool or kindergarten when they just start to learn letters and words. Several previous behavioral studies on younger children aged 4- to 5-years-old showed a refinement effect of letter knowledge on phonological awareness (Burgess and Lonigan, 1998; Lerner and Lonigan, 2016; Boets et al., 2010). In support of this early refinement effect, word reading appears to refine phonological awareness only early but not later in the first grade (Perfetti et al., 1987). Whether or not the refinement effect of reading skill on phonological representations in STG appears in emergent readers remains to be examined.

Taking a closer look at our refinement finding, we found that although all children showed an increase over development in the amount of activation in IFG, children with higher initial reading skill showed less increase. According to the neurocognitive model of language development by Skeide and Friederici (2016), the frontal lobe matures gradually and later than the temporal lobe. Studies have also shown increased activation in the dorsal IFG with age during phonological processing (e.g., Bitan et al., 2007). Thus, the observation in the current study of a brain activation increase in IFG in children from 7.5 to 9 years old likely suggests a gradual maturation of the frontal cortex. This activation increase in IFG over time was only observed when using individualized ROIs but not when using the contrast of T2 > T1 at the group level analysis, indicating that the location of phonologically sensitive voxels varied among children. Previous studies have also found that activation in the dorsal IFG increases in adult readers when a phonological task becomes more challenging, such as when segmenting phonemes, processing ambiguous speech or articulating phonologically dissimilar words (e.g., Burton et al., 2000; Okada et al., 2018; Xie and Myers, 2018). Thus, greater activation in IFG during phonological tasks could also be indicative of greater effort. The fact that we found a negative correlation between earlier reading and later IFG activation suggests that the initial higher-skilled readers at 7.5 years old likely exerted less effort in accessing phonemes during our phonological task 1.5 years later. This is consistent with our behavioral finding using standardized testing that higher reading skill predicted better phonological awareness performance.

In addition to examining brain activation within the opercular part of the left IFG as an index of phonological access, we also evaluated functional connectivity of the opercular part of IFG with the posterior left STG. However, we did not show a refinement effect of early reading skill on later functional connectivity for either onset or rhyme processing. This discrepancy between the findings for brain activation versus connectivity suggests these two measures tap into different processes (e.g., Gerchen and Kirsch, 2017). Brain activation in

IFG may indicate the effort (e.g., Alain et al., 2018; Pützer et al., 2019) of phonological access, whereas functional connectivity of IFG with STG may reflect the effectiveness of phonological access (Boets et al., 2013). An additional analysis supports this hypothesis by showing that brain activation in the IFG for onset > perceptual at T2 was negatively correlated with functional connectivity of IFG with STG for onset > perceptual at T2 [$r(59) = -.347$, $p = .007$]. Thus, our finding suggests that learning to read mainly affects the amount of effort children must use to access phonology rather than the effectiveness of accessing phonological representations.

A novelty of our study is that we distinguished different grain sizes of phonological processing, from small grain onset processing at the phonemic level to large grain processing at the rhyme level. We found that the reading skill refines later brain activation in IFG only for onset but not rhyme processing. This is consistent with the refinement hypothesis by Ziegler and Goswami (2005) that learning to read aids in the discovery of phonemes. Many previous behavioral studies have shown that reading skill is more strongly related to phonemic than rhyme awareness (see meta-analysis (Melby-Lervåg et al., 2012). These correlational studies do not provide information about the directionality of the relation, so one must look to longitudinal studies. However, the previous longitudinal studies either only examined the scaffolding effect of different grain sizes of phonological awareness on later reading skill (e.g., Muter et al., 1998; Hulme et al., 2002; Muter et al., 2004; Castles and Coltheart, 2004), or they did not examine how reading skill refines different grain sizes of phonological awareness (e.g., Wagner et al., 1997; Perfetti et al., 1987). Our neural results provide evidence that the refinement effect only occurs on small grain processing in 7.5- to 9-year-old children. A parallel behavioral analysis using in-scanner task performance for onset and rhyme judgement did not show refinement effects, suggesting that brain measures are more sensitive in detecting individual differences, consistent with some previous neural studies (e.g., Wang et al., 2020; Maurer et al., 2009). Behavioral measures are a product of many phases of processing, including cognitive control, and therefore may be less sensitive to aspects of phonological processing.

Although we did not find a refinement effect for functional connectivity, we did observe that functional connectivity of IFG with STG for rhyme processing in 7.5-year-old children predicted their reading skill one and a half years later. This scaffolding effect in older elementary school children is consistent with the only study examining the neural scaffolding effect which showed that phonological representations in STG in kindergarteners predicted children's reading skill in 5th grade (Maurer et al, 2009). However, this study did not control for initial reading, so they could not rule out the autoregressive effect. Unlike the Maurer et al. (2009) study, we controlled for the autoregressive effect and found that accessing rhyme representations in 7.5-year-old children scaffolded their reading skill 1.5 years later, providing more compelling neural evidence for the scaffolding hypothesis. In addition, our study examined different grain sizes of phonological awareness and showed that this scaffolding effect only occurred for phonological access for larger grain phonology (i.e., rhyme). This is consistent with the theory of reading development by Frith (1985) that argues for a progression to larger grain size orthography-to-phonology mapping in older skilled readers.

In contrast to the scaffolding effect found using the functional connectivity measure, we did not find such an effect using brain activation measures. As mentioned above, brain activation and functional connectivity may reflect different mechanisms, with the former one indicating effort whereas the latter one tapping into effectiveness of phonological access. In order to provide light on this, we examined the relation between brain activation in IFG for rhyme > perceptual at T1 and functional connectivity of IFG and STG for rhyme > perceptual at T1, and again we found that they were negatively correlated [$r(59) = -.261$, $p = .046$]. Thus, the lack of a scaffolding effect using brain activation could be due to the low effort that children need to access phonological representations in an easy rhyme judgement task. However, the effectiveness of accessing rhyme representations could still serve as a foundation for later efficient reading, which increasingly relies on larger grain orthography-to-phonology mapping (Frith, 1985). As with brain activation, we did not find a scaffolding effect using behavioral measures. This lack of a scaffolding effect with behavioral measures is consistent with a few previous behavioral studies (Hogan et al., 2005; Boets et al., 2010), which showed that phonological awareness in 1st or 2nd grade no longer predicted reading skill in 3rd or 4th grade. Although some previous behavioral studies found a significant scaffolding effect of phonological awareness on later reading skill up to 4th or 5th grade (e.g., Torgesen et al., 1997; Wagner et al., 1997), the effect was small (4% of variance). In fact, the scaffolding effect of early phonological awareness on later reading skill declines with development from 23% to 4% in children from kindergarten to 2nd grade and in children from 2nd to 4th grade (e.g., Wagner et al., 1997).

In conclusion, the current study examined the bidirectional relation between reading skill and phonology in the brain in children we longitudinally followed from 7.5 to 9 years old. We found that early reading skill predicted later brain activation during phonological processing using a univariate analysis. This effect was specific to onset processing in the opercular part of IFG, suggesting that reading only refines later effort in phonemic access, in alignment with the idea that learning to read helps the discovery of phonemes (Ziegler and Goswami, 2005). In a general psychophysiological interaction (gPPI) analysis, we also found that functional connectivity of IFG with STG for rhyme processing was predictive of later reading skill, providing support for the scaffolding hypothesis. This effect was specific to rhyme rather than onset, suggesting that the effectiveness of accessing larger grain phonological representations is crucial for reading acquisition in older children. This is in agreement with the argument that that older skilled readers rely more on larger grain orthography-to-phonology mappings (Frith, 1985). Overall, our findings suggest a reciprocal relation between reading and phonological access in older elementary school children. However, because both of our major findings in the current study did not survive multiple comparison correction, replication is needed.

Acknowledgements

This study was supported by the National Institutes of Health [grant numbers R01 DC013274] to James R. Booth.

Data and Code availability statement

We have made the entire project data available (Wang et al., 2021) on OpenNeuro.org. The code for imaging data analysis and the participants used for the current study are available on GitHub https://github.com/wangjinvandy/Refinement_7_9

References

Alain C, Du Y, Bernstein LJ, Barten T, Banai K, 2018. Listening under difficult conditions: an activation likelihood estimation meta-analysis. Hum. Brain Mapp. 39 (7), 2695–2709. [PubMed: 29536592]

Anthony JL, Francis DJ, 2005. Development of phonological awareness.. Curr. Dir. Psychol. Sci. 14 (5), 255–259.

Balota DA, Yap MJ, Cortese MJ, Hutchison KA, Kessler B, Loftis B, Neely JH, Nelson DL, Simpson GB, Treiman R, 2007. The English lexicon project. Behav. Res. Methods39, 445–459. [PubMed: 17958156]

Bitan T, Cheon J, Lu D, Burman DD, Gitelman DR, Mesulam MM, Booth JR, 2007. Developmental changes in activation and effective connectivity in phonological processing. Neuroimage38 (3), 564–575. [PubMed: 17884585]

Boets B, 2014. Dyslexia: reconciling controversies within an integrative developmental perspective. Trends Cognit. Sci. 18 (10), 501–503. [PubMed: 25034040]

Boets B, De Smedt B, Cleuren L, Vandewalle E, Wouters J, Ghesquiere P, 2010. Towards a further characterization of phonological and literacy problems in Dutch-speaking children with dyslexia. Br. J. Dev. Psychol. 28 (1), 5–31. [PubMed: 20306623]

Boets B, Op de Beeck HP, Vandermosten M, Scott SK, Gillebert CR, Mantini D, Bulthe J, Sunaert S, Wouters J, Ghesquiere P, 2013. Intact but less accessible phonetic representations in adults with dyslexia. Science342 (6163), 1251–1254. [PubMed: 24311693]

Bolger DJ, Hornickel J, Cone NE, Burman DD, Booth JR, 2008. Neural correlates of orthographic and phonological consistency effects in children. Hum. Brain Mapp. 29 (12), 1416–1429. [PubMed: 17957704]

Brennan C, Cao F, Pedroarena-Leal N, McNorgan C, Booth JR, 2013. Reading acquisition reorganizes the phonological awareness network only in alphabetic writing systems. Hum. Brain Mapp. 34 (12), 3354–3368. [PubMed: 22815229]

Burgess SR, Lonigan CJ, 1998. Bidirectional relations of phonological sensitivity and prereading abilities: evidence from a preschool sample. J. Exp. Child. Psychol. 70 (2), 117–141. [PubMed: 9729452]

Burton MW, Small SL, Blumstein SE, 2000. The role of segmentation in phonological processing: an fMRI investigation. Journal of cognitive neuroscience12 (4), 679–690. [PubMed: 10936919]

Castles A, Coltheart M, 2004. Is there a causal link from phonological awareness to success in learning to read?Cognition91 (1), 77–111. [PubMed: 14711492]

Chall JS, 1983. Stages of Reading Development. McGraw-Hill, New York.

Corina DP, Richards TL, Serafini S, Richards AL, Steury K, Abbott RD, Echelard DR, Maravilla KR, Berninger VW, 2001. fMRI auditory language differences between dyslexic and able reading children. Neuroreport12 (6), 1195–1201. [PubMed: 11338191]

Dębska A, Chyl K, Dziegiel G, Kacprzak A, Luniewska M, Plewko J, Marchewka A, Grabowska A, Jednoróg K, 2019. Reading and spelling skills are differentially related to phonological processing: behavioral and fMRI study. Dev. Cognit. Neurosci. 39.

Dębska A, Luniewska M, Chyl K, Banaszkiewicz A, Ż elechowska A, Wypych M, Marchewka A, Pugh KR, Jednoróg K, 2016. Neural basis of phonological awareness in beginning readers with familial risk of dyslexia—results from shallow orthography. Neuroimage132, 406–416. [PubMed: 26931814]

Dixon WJ, 1960. Simplified estimation from censored normal samples. Ann. Math. Stat. 3, 385–391.

Eklund A, Nichols TE, Knutsson H, 2016. Cluster failure: why fMRI inferences for spatial extent have inflated false-positive rates. Proc. Natl. Acad. Sci. 113 (28), 7900–7905. [PubMed: 27357684]

Frith U, 1985. Beneath the surface of developmental dyslexia. Surf. Dyslexia32, 301–330.

Gerchen MF, Kirsch P, 2017. Combining task-related activation and connectivity analysis of fMRI data reveals complex modulation of brain networks. Hum. Brain Mapp. 38 (11), 5726–5739. [PubMed: 28782871]

Hagoort P, 2014. Nodes and networks in the neural architecture for language: broca's region and beyond. Curr. Opin. Neurobiol. 28, 136–141. [PubMed: 25062474]

Hickock G, Poeppel D, 2000. Towards a functional neuroanatomy of speech perception. Trends Cognit. Sci. 4 (4), 131–138. [PubMed: 10740277]

Hogan TP, Catts HW, Little TD, 2005. The relationship between phonological awareness and reading. Lang. Speech Hear. Serv. Sch. 36, 285–293. [PubMed: 16389701]

Hulme C, Hatcher PJ, Nation K, Brown A, Adams J, Stuart G, 2002. Phoneme awareness is a better predictor of early reading skill than onset-rime awareness. J. Exp. Child. Psychol. 82 (1), 2–28. [PubMed: 12081455]

Kaufman AS, & Kaufman NL (2004). Kaufman brief intelligence test, Second Edition. Bloomington, MN: Pearson.Kovelman I, Norton ES, Christodoulou JA, Gaab N, Lieberman DA, Triantafyllou C, Wolf M, Whitefield-Gabrieli S, & Gabrieli JD (2012). Brain basis of phonological awareness for spoken language in children and its disruption in dyslexia. Cerebral Cortex, 22(4), 754–764. [PubMed: 21693783]

Kovelman I, Norton ES, Christodoulou JA, Gaab N, Lieberman DA, Triantafyllou C, Gabrieli JD, 2012. Brain basis of phonological awareness for spoken language in children and its disruption in dyslexia. Cerebral Cortex22 (4), 754–764. [PubMed: 21693783]

Kraft I, Schreiber J, Cafiero R, Metere R, Schaadt G, Brauer J, Neef NF, Muller B, Kristen H, Wilke A, Boltze J, Friederici AD, Skeide MA, 2016. Predicting early signs of dyslexia at a preliterate age by combining behavioral assessment with structural MRI. Neuroimage143, 378–386. [PubMed: 27608602]

Kuhl U, Neef NE, Kraft I, Schaadt G, Dörr L, Brauer J, Czepezauer I, Muller B, Wilcke A, Kirsten H, Emmrich F, Boltze J, Friederici AD, Skeide AS, 2020. The emergence of dyslexia in the developing brain. Neuroimage211, 116633. [PubMed: 32061802]

Leonard MK, Chang EF, 2014. Dynamic speech representations in the human temporal lobe. Trends Cognit. Sci. 18 (9), 472–479. [PubMed: 24906217]

Lerner MD, Lonigan CJ, 2016. Bidirectional relations between phonological awareness and letter knowledge in preschool revisited: a growth curve analysis of the relation between two code-related skills. J. Exp. Child. Psychol. 144, 166–183. [PubMed: 26745710]

Lundberg I, Frost J, Petersen O-P, 1988. Effects of an extensive program for stimulating phonological awareness in preschool children. Read. Res. Q. 23 (3), 263–284.

Maurer U, Bucher K, Brem S, Benz R, Kranz F, Schulz E, van der Mark S, Steinhausen H-C, Brandeis D, 2009. Neurophysiology in preschool improves behavioral prediction of reading ability throughout primary school. Biol. Psychiatry66 (4), 341–348. [PubMed: 19423082]

Mazaika P, Hoeft F, Glover GH, Reiss AL, 2009. Methods and software for fMRI analysis for clinical subjects. Human Brain Mapping Conference.

McLaren DG, Ries ML, Xu G, Johnson SC, 2012. A generalized form of context-dependent psychophysiological interactions (gPPI): a comparison to standard approaches. Neuroimage61 (4), 1277–1286. [PubMed: 22484411]

Melby-Lervåg M, Lyster SAH, Hulme C, 2012. Phonological skills and their role in learning to read: a meta-analytic review. Psychol. Bull. 138 (2), 322–352. [PubMed: 22250824]

Muter V, Hulme C, Snowling M, 1998. Segmentation, not rhyming, predicts early progress in learning to read. J. Exp. Child. Psychol. 71, 3–27. [PubMed: 9742182]

Muter V, Hulme C, Snowling MJ, Stevenson J, 2004. Phonemes, rimes, vocabulary, and grammatical skills as foundations of early reading development: evidence from a longitudinal study. Dev. Psychol. 40 (5), 665–681. [PubMed: 15355157]

Myers EB, Blumstein SE, Walsh E, Eliassen J, 2009. Inferior frontal regions underlie the perception of phonetic category invariance. Psychol. Sci. 20 (7), 895–903. [PubMed: 19515116]

Nelson DL, McEvoy CL, & Schreiber TA (1998). The University of South Florida word association, rhyme, and word fragment norms. http://www.usf.edu/FreeAssociation/.

Okada K, Matchin W, Hickok G, 2018. Phonological feature repetition suppression in the left inferior frontal gyrus. Journal of cognitive neuroscience30 (10), 1549–1557. [PubMed: 29877763]

Perfetti CA, Beck I, Bell LC, Hughes C, 1987. Phonemic knowledge and learning to read are reciprocal: a longitudinal study of first grade children. Merrill Palmer Q. 1982-, 283–319.

Peterson RL, Pennington BF, 2015. Developmental dyslexia. Annu. Rev. Clin. Psychol. 11, 283–307. [PubMed: 25594880]

Pützer M, Moringlane JR, Sikos L, Reith W, Krick CM, 2019. fMRI and acoustic analyses reveal neural correlates of gestural complexity and articulatory effort within bilateral inferior frontal gyrus during speech production. Neuropsychologia132, 107–129.

Ramus F, 2014. Neuroimaging sheds new light on the phonological deficit in dyslexia. Trends Cognit. Sci. 18 (6), 274–275. [PubMed: 24568928]

Raschle NM, Zuk J, Gaab N, 2012. Functional characteristics of developmental dyslexia in left-hemispheric posterior brain regions predate reading onset. Proc. Natl. Acad. Sci. 109 (6), 2156–2161. [PubMed: 22308323]

Seymour HN, Roeper T, De Villiers JG, 2003. DELV: Diagnostic Evaluation of Language Variation: Screening Test. PsychCorp.

Skeide MA, Friederici AD, 2016. The ontogeny of the cortical language network. Nat. Rev. Neurosci. 17 (5), 323–332. [PubMed: 27040907]

Torgesen JK, Wagner RK, Rashotte CA, Burgess S, Hecht S, 1997. Contributions of phonological awareness and rapid automatic naming ability to the growth of word-reading skills in second-to fifth-grade children. Scientific studies of reading1 (2), 161–185.

Treiman R, Zukowski A, 1991. Levels of phonological awareness. Phonological Processes in Literacy: a Tribute to. Isabelle Y. Liberman, pp. 67–83.

Vandermosten M, Correia J, Vanderauwera J, Wouters J, Ghesquière P, Bonte M, 2019. Brain activity patterns of phonemic representations are atypical in beginning readers with family risk for dyslexia. Dev. Sci. doi: 10.1111/desc.12857.

Vitevitch MS, Luce PA, 2004. A web-based interface to calculate phonotactic probability for words and nonwords in English. Behav. Res. Methods. Instrum. Comput. 36, 481–487. [PubMed: 15641436]

Wagner RK, Torgesen JK, Rashotte CA, Hecht SA, Barker TA, Burgess SR, Donahue J, Garon T, 1997. Changing relations between phonological processing abilities and word-level reading as children develop from beginning to skilled readers: a 5-year longitudinal study. Dev. Psychol. 33 (3), 468–479. [PubMed: 9149925]

Wang J, Joanisse MF, Booth JR, 2018. Reading skill related to left ventral occipitotemporal cortex during a phonological awareness task in 5–6-year-old children. Dev. Cognit. Neurosci. 30, 116–122. [PubMed: 29454219]

Wang J, Joanisse MF, Booth JR, 2020. Neural representations of phonology in temporal cortex scaffold longitudinal reading gains in 5- to 7-year-old children. Neuroimage207Article 116359.

Wang J, Joanisse MF, Booth JR, 2021. Letter fluency in 7–8-year-old children is related to the anterior, but not posterior, ventral occipito-temporal cortex during an auditory phonological task. Dev. Cognit. Neurosci. 47, 100898. [PubMed: 33341533]

Wiig EH, Semel E, Secord WA, 2013. Clinical Evaluation Language Fundamentals (CELF-5), fifth ed. Pearson, San Antonio.

Wilke M, Altaye M, Holland SK, Authorship Consortium CMIND, 2017. CerebroMatic: a versatile toolbox for spline-based MRI template creation. Front. Comput. Neurosci. 11, 5. [PubMed: 28275348]

Woodcock RW, McGrew KS, Mather N, 2001. Woodcock-Johnson III. Riverside Publishing, Itasca, IL.

Xie X, Myers E, 2018. Left inferior frontal gyrus sensitivity to phonetic competition in receptive language processing: A comparison of clear and conversational speech. Journal of cognitive neuroscience30 (3), 267–280. [PubMed: 29160743]

Ziegler JC, Goswami U, 2005. Reading acquisition, developmental dyslexia, and skilled reading across languages: a psycholinguistic grain size theory. Psychol. Bull. 131 (1), 3–29. [PubMed: 15631549]

Wagner R, Torgesen J, Rashotte C & Pearson NA (2013). Comprehensive Test of Phonological Processing-2nd ed. (CTOPP-2). Pro-ed.

Wang J, Lytle MN, Weiss Y, Yamasaki BL, & Booth JR (2021, 4 2). A longitudinal neuroimaging dataset on language processing in children ages 5, 7, and 9 years old. 10.31234/osf.io/tpndf.

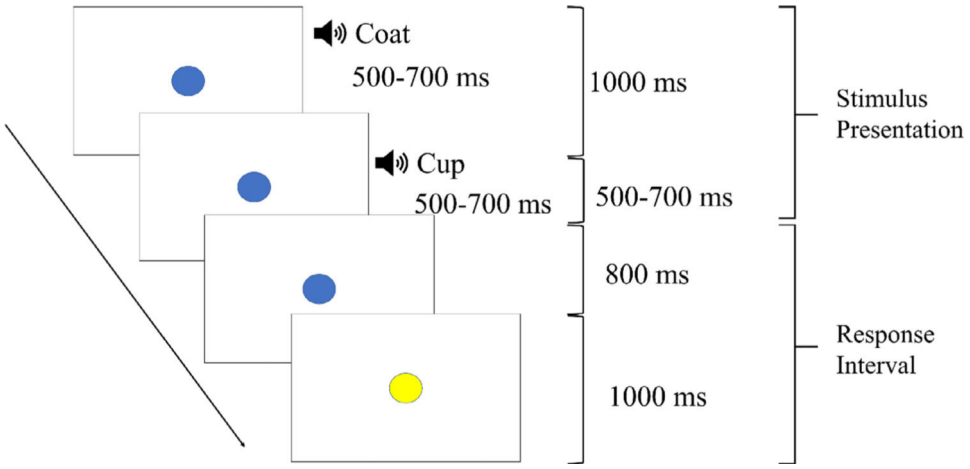

Fig. 1.
Procedure for the auditory phonological awareness task.

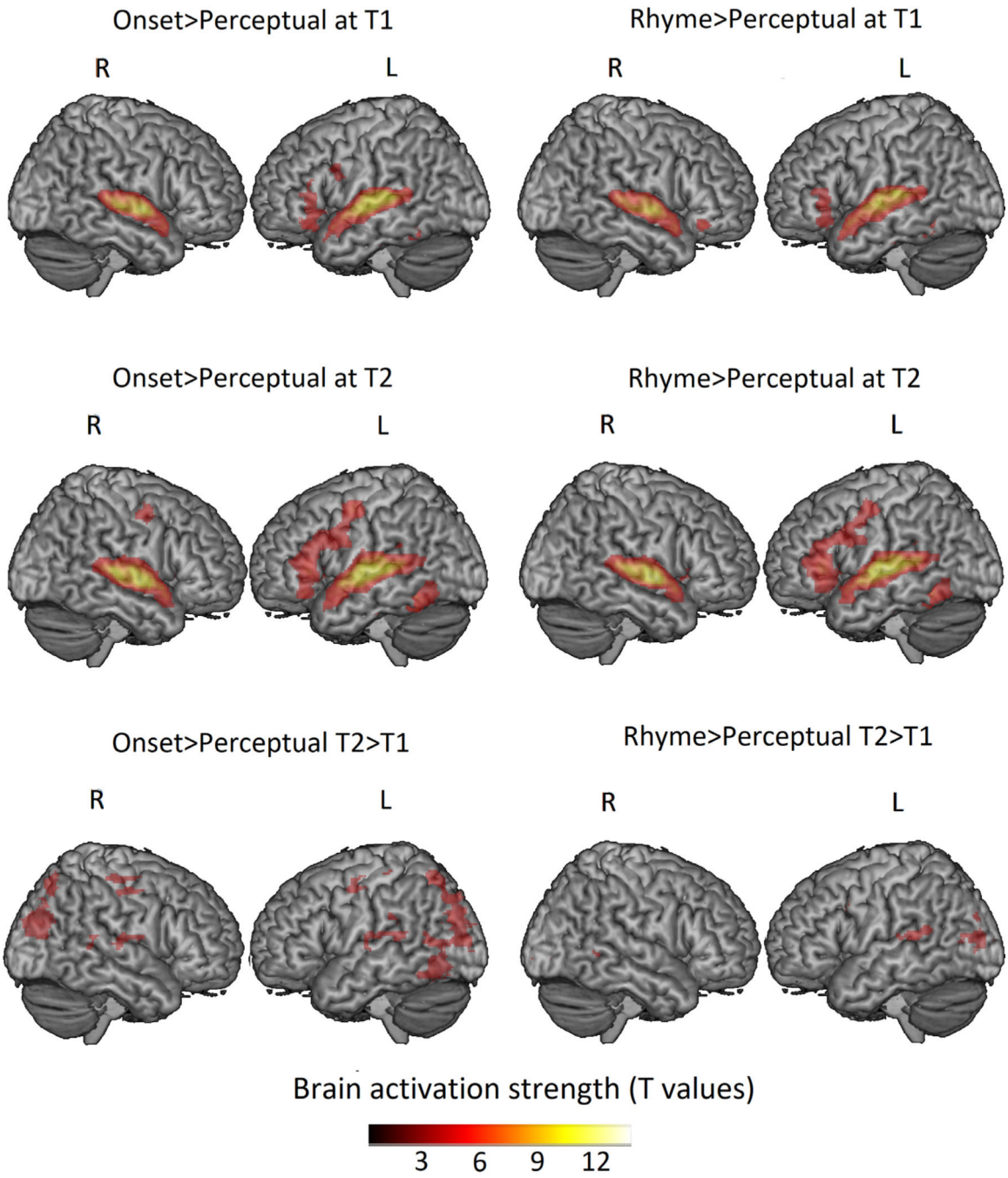

Fig. 2.

Group level brain activation during onset and rhyme processing at both T1 and T2 and T2>T1. Group maps thresholded at voxel-wise $p < 0.001$ uncorrected and cluster-wise $p < 0.05$ corrected within the whole brain mask. Clusters with size greater than 88 voxels are shown. L = left hemisphere; R = right hemisphere.

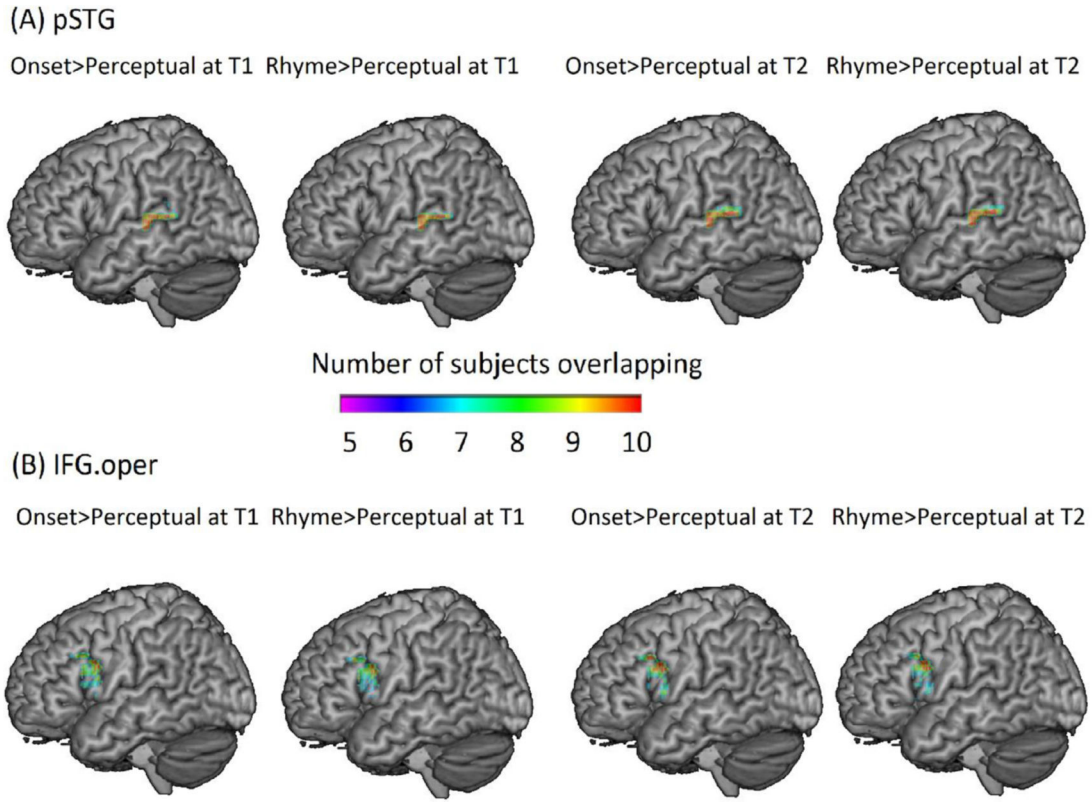

Fig. 3.

Regions of interest in temporal and frontal cortex. (A) Overlap of individualized ROI in the posterior superior temporal gyrus (pSTG) (B) Overlap of individualized ROI in the opercular part of inferior frontal gyrus (IFG.oper). The ROIs defined at T1 (in the left panel) were used in the examination of the scaffolding hypothesis. The ROIs defined at T2 (in the right panel) were used in the examination of the refinement hypothesis.

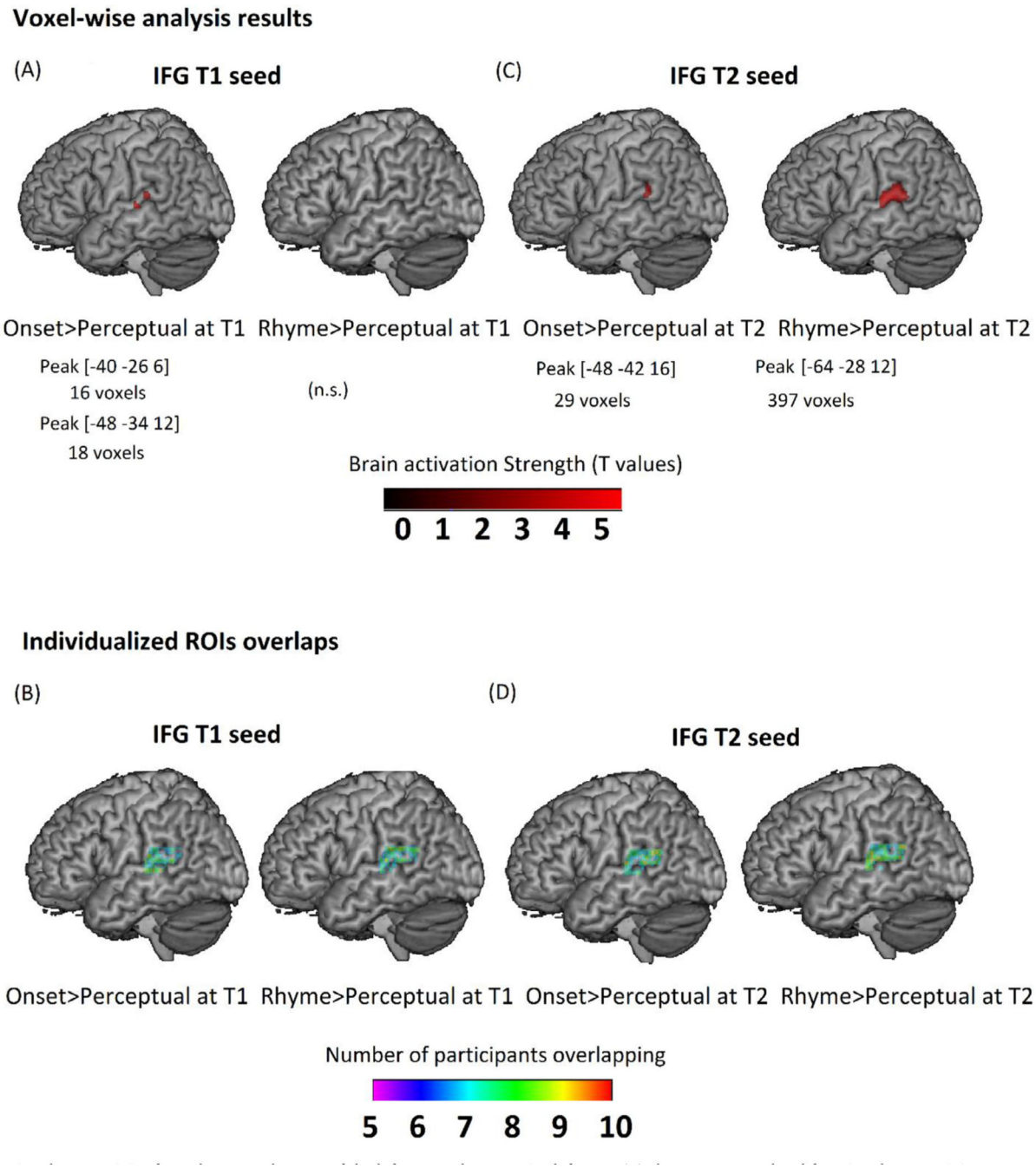

Fig. 4.

Functional connectivity from the opercular part of the left IFG to the posterior left STG. (A) shows T1 group-level functional connectivity maps using the top 100 most activated voxels in IFG for either onset > perceptual or rhyme > perceptual at T1 as the seed region. Group maps are thresholded at voxel-wise $p < 0.001$ (T value > 3.23) uncorrected and cluster-wise $p < 0.05$ corrected within the posterior left STG mask. Clusters greater than 9 voxels are shown. Peak coordinates and cluster sizes are reported in the figure. (B) shows the overlap of individualized functional connectivity regions of interest (ROIs) for onset >

perceptual or rhyme > perceptual within the posterior left STG at T1. These individualized ROIs were used to examine the scaffolding hypothesis. (C) shows T2 group-level functional connectivity maps using the top 100 most activated voxels in IFG for either onset > perceptual or rhyme > perceptual at T2 as the seed region. Group maps thresholded at voxel-wise $p < 0.001$ uncorrected (T value > 3.23) and cluster-wise $p < 0.05$ corrected within the posterior left STG mask. Clusters greater than 9 voxels are shown. Peak coordinates and cluster sizes are reported in the figure. (D) shows the overlap of individualized functional connectivity regions of interest for onset > perceptual or rhyme > perceptual within the posterior left STG at T2. These individualized ROIs were used to examine the refinement hypothesis.

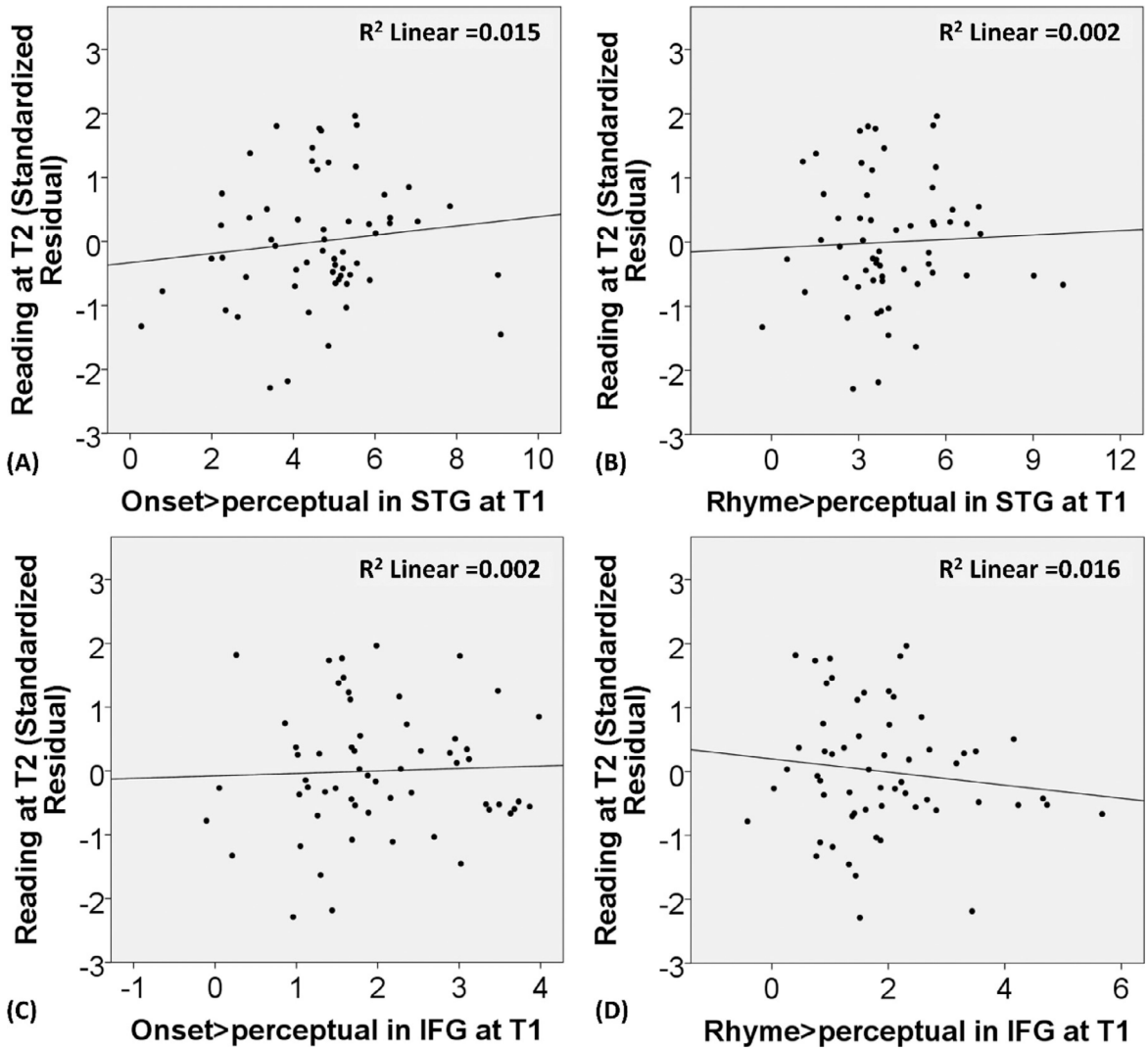

Fig. 5.

The scatterplots for the relation between brain activation in the posterior left STG and the opercular part of the left IFG for onset > perceptual and rhyme > perceptual at T1 and the standardized residuals of reading skill at T2 after controlling for reading skill and non-verbal IQ at T1.

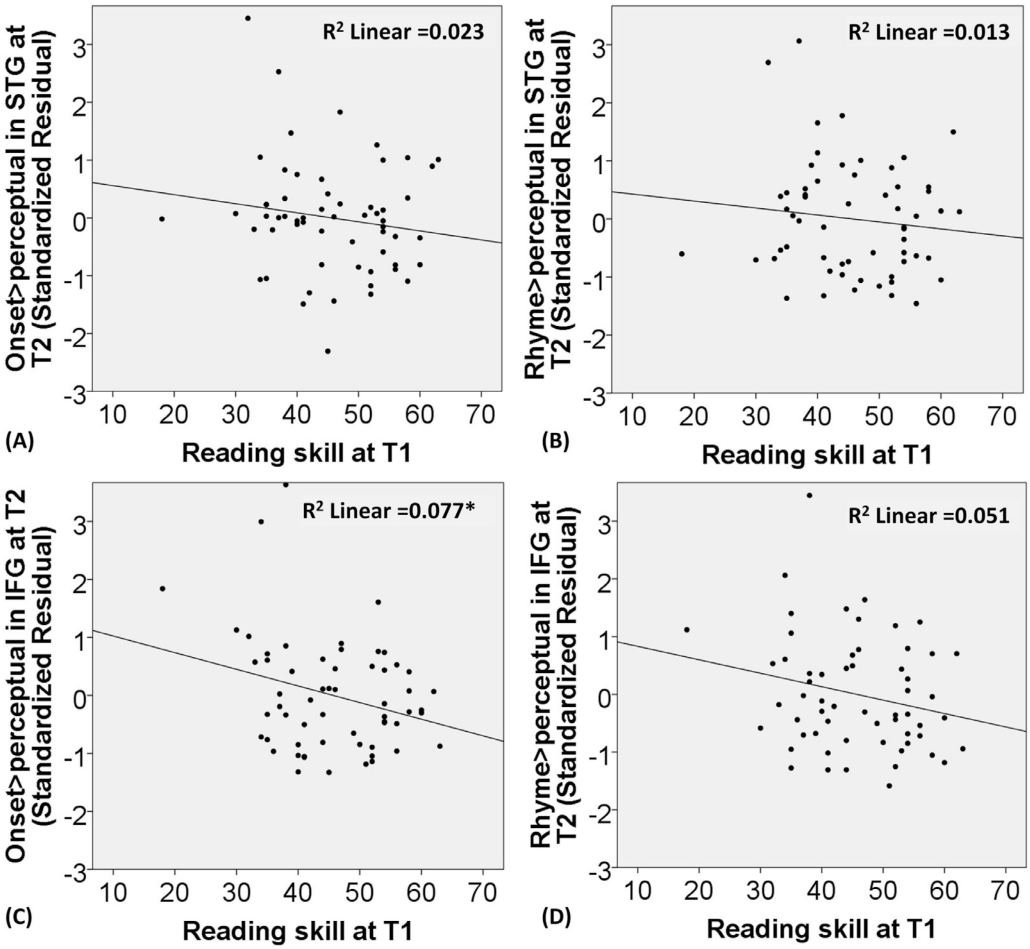

Fig. 6.

The scatterplots for the relation between reading skill at T1 (raw score) and the standardized residuals of brain activation in the posterior left STG and the opercular part of the left IFG for onset > perceptual and rhyme > perceptual at T2 after controlling for brain activation and non-verbal IQ at T1. * indicates $p < 0.05$ uncorrected.

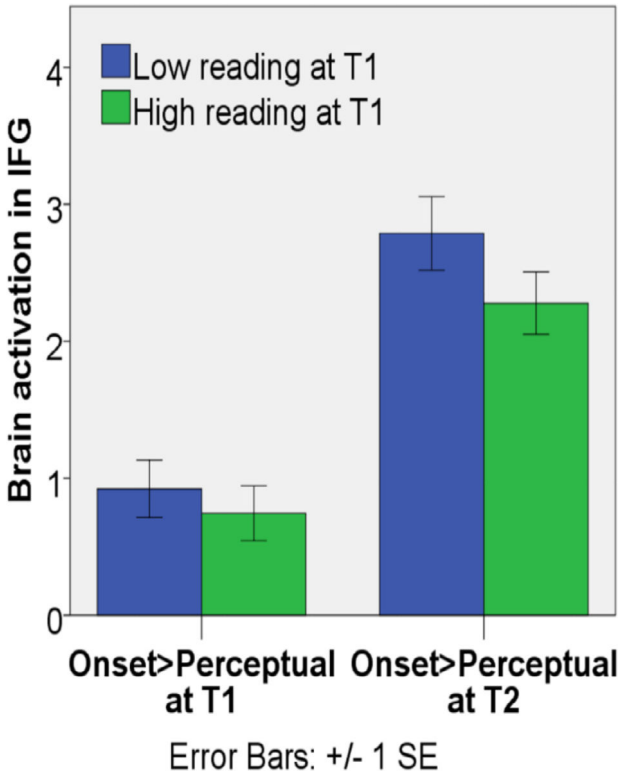

Fig. 7.

Brain activation in the opercular part of the left IFG for onset>perceptual at T1 and T2 for high (green) and low (blue) T1 reading groups based on a median split.

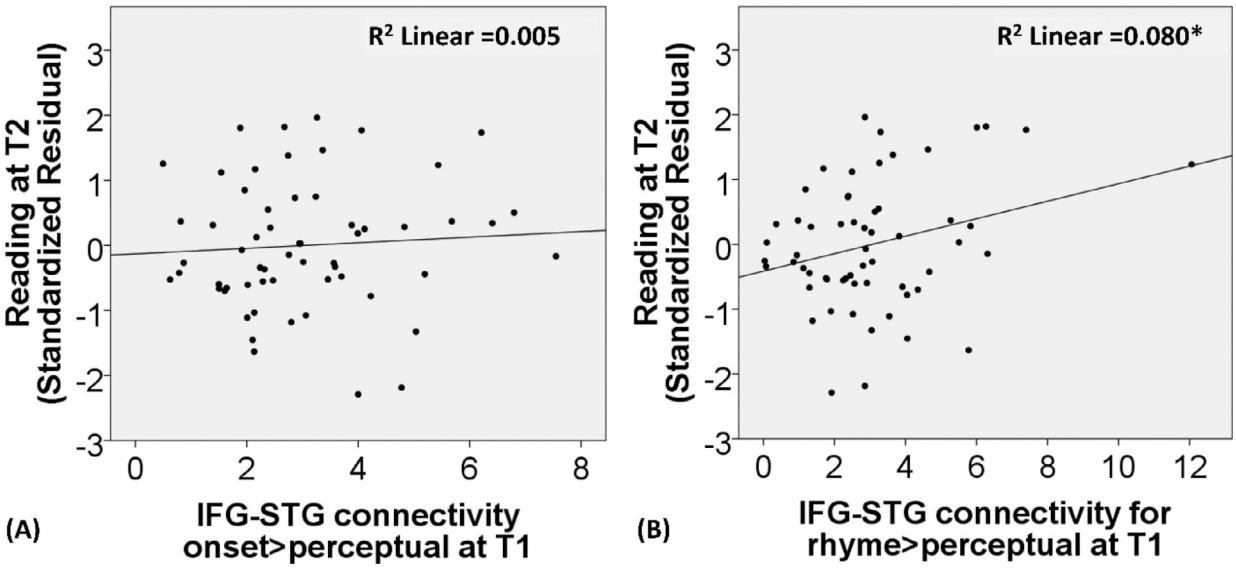

Fig. 8.

The scatterplots for the relation between functional connectivity of the opercular part of the left IFG with the posterior left STG for onset > perceptual and rhyme > perceptual at T1 and the standardized residuals of reading skill at T2 after controlling for reading skill and non-verbal IQ at T1. * indicates $p < 0.05$ uncorrected.

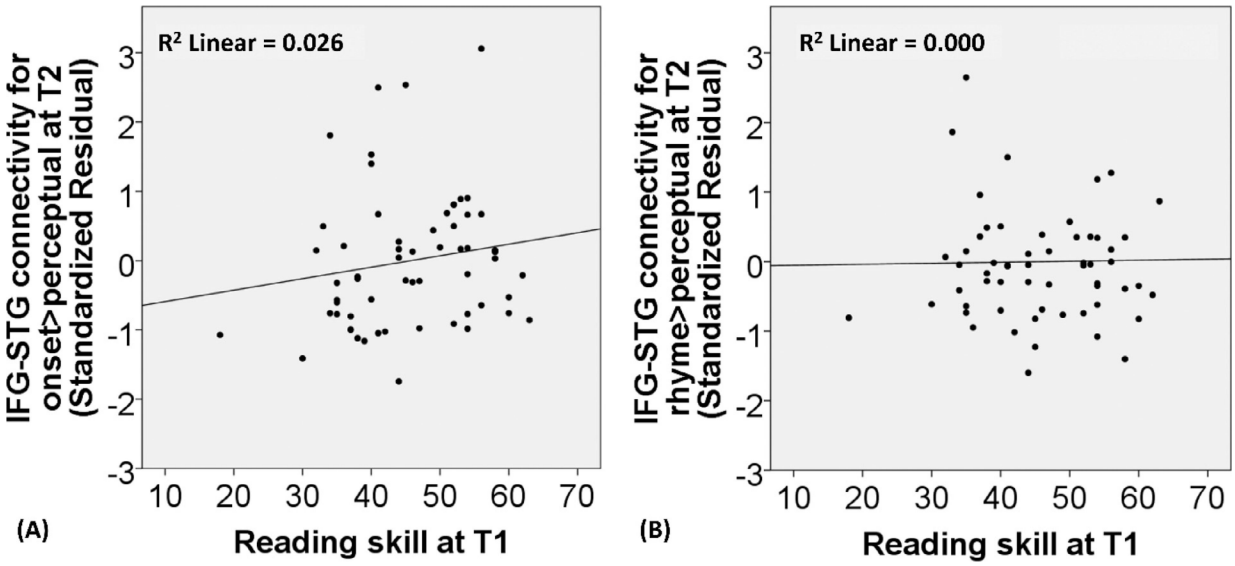

Fig. 9.

The scatterplots for the relation between reading skill at T1 (raw score) and the standardized residuals of brain connectivity of the opercular part of the left IFG with the posterior left STG for onset > perceptual and rhyme > perceptual at T2 after controlling for brain connectivity and non-verbal IQ at T1.

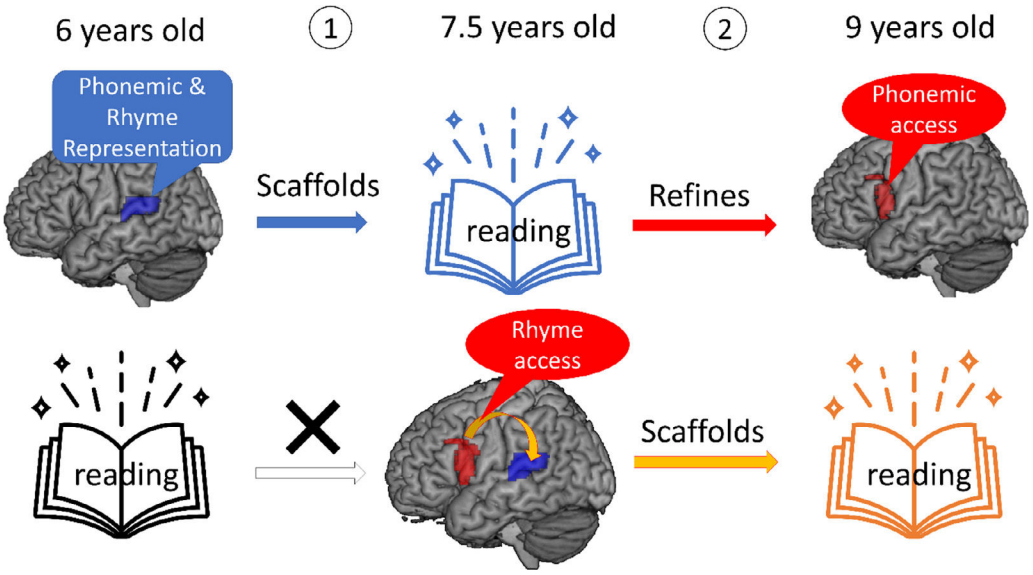

Fig. 10.

The developmental progression in the relation between phonological processing in the brain and reading ability. (1) is supported by our previous study (Wang et al., 2021) and (2) is supported by the current study.

Table 1

Descriptive statistics of demographics and standardized testing scores.

		Number of Females	Number of Males		Mean of Age (SD)	
Demographics	T1	32	27		7.33 (0.28)	
	T2	32	27		9.17 (0.15)	
			Raw score		Standardized score	
			Mean (SD)	Range	Mean (SD)	Range
Screening/Control Variables	T1	CELF-5 core language	44.6 (8.7)	28–62	107.5 (14.5)	82–137
		KBIT-2 non-verbal IQ	26.3 (6.8)	12–41	110.0 (17.6)	74–147
Variables of Interest	T1	CTOPP-2 phonological awareness	33.2 (6.5)	19–48	107.7 (14.2)	77–140
		WJ-III letter word identification	45.7 (9.5)	18–63	118.7 (12.3)	83–138
	T2	CTOPP-2 phonological awareness	30.2 (5.6)	20–44	101.2 (12.1)	80–131
		WJ-III letter word identification	56.0 (7.7)	31–69	113.5 (12.1)	76–133

Table 2

Examples of the stimuli in the auditory phonological judgment task.

Condition	Response	Brief Explanation	Example
Onset	Yes	The two words start with the same sound	Coat – Cup
Rhyme	Yes	The two words rhyme	Wide – Ride
Non-match	No	The two words have no same sounds	Zip – Cone
Perceptual	Yes	Frequency modulated noise	“Shh – Shh”

Table 3

Accuracies for different conditions during the auditory phonological task.

Condition	Time 1 (%) (Mean±SD)	[range]	Time 2 (%) (Mean±SD)	[range]
Onset	(69.3 ± 15.1)	[25.0–91.7]	(78.2 ± 15.2)	[29.2–100]
Rhyme	(88.5 ± 9.7)	[58.3–100]	(93.4 ± 7.0)	[70.8–100]
Non-Match	(81.4 ± 12.5)	[45.8–100]	(89.5 ± 7.5)	[66.7–100]
Perceptual	(93.7 ± 8.1)	[66.7–100]	(97.4 ± 3.9)	[83.3–100]

Table 4

Group level brain activation for onset and rhyme processing at both T1 and T2.

Contrast	Brain region	Brodmann Area	Coordinate (peak)	Voxel (2mm)	T
Onset > Perceptual at T1					
	Left superior temporal gyrus	22	−68 –26 8	2246	14.08
	Right superior temporal gyrus	22	64 –4 −2	1770	14.07
	Left inferior temporal gyrus	20/37	−46 –48 −20	406	7.04
	Left inferior frontal gyrus	47/45	−38 30 –14	845	5.60
	Left precentral gyrus	9	−48 10 30	128	4.15
Rhyme > Perceptual at T1					
	Left superior temporal gyrus	22	−66 –26 8	2255	12.48
	Right superior temporal gyrus	22	66 –6 −2	1785	12.37
	Left fusiform gyrus	37	−42 –42 −22	442	6.79
	Right inferior frontal gyrus	47	40 32 –14	132	6.33
	Left inferior frontal gyrus	47/45	−48 28 0	438	5.59
Onset > Perceptual at T2					
	Left superior temporal gyrus	22	−62 –12 4	3275	14.98
	Right superior temporal gyrus	22	62 –8 2	2778	13.31
	Left fusiform gyrus	37	−42 –46 −18	1359	8.28
	Left anterior cingulum	24	−6 4 28	129	8.24
	Left supplementary motor area	6	−4 6 60	405	7.13
	Left precentral gyrus /Left inferior frontal gyrus	6/9/46/45	−54 –6 52	2795	6.89
	Left caudate	-	−14 8 6	907	6.68
	Right fusiform area	37	42 –40 −16	108	6.33
	Left hippocampus	-	−20 –18 −18	91	6.17
	Right fusiform gyrus	37	22 –36 −16	146	6.08
	Right precentral gyrus	6	54 –6 46	145	5.33
	Right insula	13	32 30 2	110	4.65
Rhyme > Perceptual at T2					
	Left superior temporal gyrus/ Left inferior frontal gyrus	22/9/46	−64 –12 4	5358	15.13
	Right superior temporal gyrus	22	62 –6 −2	2756	13.90

Contrast	Brain region	Brodmann Area	Coordinate (peak)	Voxel (2mm)	T
	Left fusiform gyrus	37	−46 –48 −16	946	8.57
	Right fusiform gyrus	37	42 –38 −16	90	6.33
	Left pallidum	-	−16 6 –2	405	5.55
	Right insula	13	36 16 2	103	5.23
	Left hippocampus	-	−20 –16 −18	120	4.93
Onset > Perceptual for T2>T1					
	Left middle occipital/inferior parietal lobule	19/7	−30 –70 28	3353	5.91
	Right superior occipital gyrus lobule	19/7	28 –68 28	1650	5.22
	Right insular	13	28 –22 22	176	5.09
	Left postcentral gyrus	4	−22 –30 62	401	4.98
	Left supramarginal gyrus/inferior parietal lobule	40	−44 –34 28	131	4.49
	Right postcentral gyrus	4	60 –12 48	336	4.73
	Right fusiform gyrus	19/37	20 –62 −14	235	4.71
	Right postcentral gyrus	4	50 –22 58	199	4.71
	Left supplementary motor area	6	−2 –14 54	375	4.49
	Left superior temporal gyrus	22/42	−66 –38 14	194	4.44
	Right superior temporal gyrus	42	62 –20 −12	148	4.39
	Right superior temporal gyrus	22	58 –40 14	90	3.94
Rhyme > Perceptual for T2 > T1					
	Left superior temporal gyrus	22/42	−62 –40 12	361	5.37
	Left cerebellum	-	−4 –56 −6	88	4.94
	Left Calcarine	30/18	−20 –66 4	1087	4.67
	Right fusiform gyrus	19/37	34 –56 −6	116	4.54
	Left precentral gyrus	9	−32 2 34	95	4.44
	Right lingual gyrus	18	16 –80 −6	143	4.42
	Left middle occipital gyrus	18	−40 –86 12	143	3.99

Table 5

The result of the hierarchical regression analyses examining the scaffolding hypothesis using brain activation.

		Dependent measure											
		Reading skill at T2											
	Predictor	β	R^2	ΔR^2	β	R^2	ΔR^2	β	R^2	ΔR^2	β	R^2	ΔR^2
Model 1	Non-verbal IQ	.002			.002			.002			.002		
	Reading skill at T1	.855 ***	.731		.855 ***	.731		.855 ***	.731		.855 ***	.731	
Model 2	Non-verbal IQ	.003			.004			.007			−.020		
	Reading skill at T1	.856 ***			.855 ***			.854 ***			.858 ***		
	Onset>Perceptual in STG at T1	.064	.735	.004									
	Rhyme> Perceptual in STG at T1				.023	.732	.001						
	Onset>Perceptual in IFG at T1							.021	.732	.001			
	Rhyme> Perceptual in IFG at T1										−.072	.736	.005

* = $p < 0.05$

** = $p < 0.01$

*** = $p < 0.001$ uncorrected.

Table 6

The result of the hierarchical regression analyses examining the refinement hypothesis using brain activation.

		Dependent measure											
		Onset>Perceptual STG at T2			**Rhyme>Perceptual STG at T2**			**Onset>Perceptual IFG at T2**			**Rhyme>Perceptual IFG at T2**		
	Predictor	*β*	R^2	ΔR^2	*β*	R^2	ΔR^2	*β*	R^2	ΔR^2	*β*	R^2	ΔR^2
Model 1	Non-verbal IQ	−.036			0.43			−.173			−.073		
	Onset >Perceptual in STG at T1	.470***	.223										
	Rhyme>Perceptual in STG at T1				.397***	.155							
	Onset >Perceptual in IFG at T1							.275*	.119				
	Rhyme>Perceptual in IFG at T1										.212	.059	
Model 2	Non-verbal IQ	−.022			.088			−.058			.029		
	Onset >Perceptual in STG at T1	.466***											
	Rhyme>Perceptual in STG at T1				.398**								
	Onset >Perceptual in IFG at T1							.295*					
	Rhyme>Perceptual in IFG at T1										.238		
	Reading skill at T1	−.155	.244	.021	−.124	.168	.013	−.303*	.198	.079	−.256	.115	.056

* = $p < 0.05$

** = $p < 0.01$

*** = $p < 0.001$ uncorrected.

Table 7

The result of the hierarchical regression analyses examining the scaffolding hypothesis using brain connectivity.

		Dependent measure					
		Reading skill at T2					
	Predictor	β	R^2	ΔR^2	β	R^2	ΔR^2
Model 1	Non-verbal IQ.	.002			.002		
	Reading skill at T1	.855 ***	.731		.855 ***	.731	
Model 2	Non-verbal IQ	.004			−.012		
	Reading skill at T1	.860 ***			.890 ***		
	IFG-STG connectivity for onset > perceptual at T1	.037	.732	.001			
	IFG-STG connectivity for rhyme > perceptual at T1				.154 *	.754	.023

* = $p < 0.05$

** = $p < 0.01$

*** = $p < 0.001$ uncorrected.

Table 8

The result of the hierarchical regression analyses examining the refinement hypothesis using functional connectivity.

		Dependent measure					
		IFG-STC connectivity for onset > perceptual at T2			IFG-STG connectivity for rhyme > perceptual at T2		
	Predictor	β	R^2	ΔR^2	β	R^2	ΔR^2
Model 1	Non-verbal IQ	−.129			−.273 *		
	IFG-STG connectivity for onset > perceptual at T1	.170	.047				
	IFG-STG connectivity for rhyme > perceptual at T1						
	IFG-STG connectivity for rhyme > perceptual at T1				−.008	.074	
Model 2	Non-verbal IQ	−.196			−.279		
	IFG-STG connectivity for onset > perceptual at T1	.174					
	IFG-STG connectivity for rhyme > perceptual at T1				−.007		
	Reading skill at T1	.181	.075	.028	.016	.074	.000

* = $p < 0.05$

** = $p < 0.01$

*** = $p < 0.001$ uncorrected.