1

The effect of a reading aloud program on reading rate and reading prosody in a group of sixth-grade low-achievement, language-minority, and/or low-SES readers

Juliette Quadri, Justine Masson, Martine Poncelet

Psychology and Neuroscience of Cognition Research Unit, University of Liège (ULiège), Place des Orateurs, 1, 4000, Liège, Belgium

Abstract: The purpose of this study was to investigate whether a group of sixth-grade low-achievement, language-minority (LM) and/or low-socioeconomic-status (low-SES) students' reading skills could be improved by an intervention designed to train reading fluency, a critical component of reading development that is defined as integrating speed and accuracy (reading rate) as well as prosody. The study included 54 sixth-grade LM/low-SES students attending a low-SES school. Two groups matched on different reading and control measures were randomly assigned to an intervention condition that consisted of a tutored reading aloud intervention combining repeated reading, assisted wide reading, guided comprehension and modeling prosody or a control condition where students completed their usual classroom activities. Reading rate (number of words read correctly per minute), prosody and comprehension were assessed. Students who received the intervention scored better on measures of reading prosody but not on reading rate and comprehension measures than students from the control group. The discussion questions the lack of apparent results for reading rate in relation to the progress made in reading prosody.

2

The Belgian school system includes the existence of schools, mostly urban, where students from low-income backgrounds are concentrated. These schools often have a high proportion of language-minority (LM) students, or "students who come from home where the primary language spoken is not the language of schooling" (Kieffer & Lesaux, 2012). These LM students and their French-speaking peers grow up in low-income communities and attend low-SES schools, factors that contribute to putting them at risk of becoming struggling readers (Buckingham et al., 2013).

Even if LM students develop proficient word decoding skills in the primary elementary grades, their reading comprehension remains less proficient than that of their native peers (L1) in the upper grades (Farnia & Geva, 2013; Geva & Farnia, 2012; Kim et al., 2012; Perfetti, 2007; Raudszus et al., 2021; Verhoeven & van Leeuwe, 2012). However, the gap between LM and L1 students decreases greatly when comparing LM students with their peers from similar economic backgrounds (Kieffer, 2011; Lesaux & Kieffer, 2010).

The successive Programs for International Student Assessment (PISAs) indicate that in the French-speaking part of Belgium where the present study took place, students with an immigrant background and those from low-income homes repeatedly obtained lower performance in reading comprehension than other students (Bricteux et al., 2021; Jacobs et al., 2009). Given that a link between reading comprehension and reading fluency has been suggested for L1 students (Fuchs et al., 2001; Jenkins et al., 2003; Rasinski et al., 2009) and for LM students (Riedel, 2007; Wiley & Deno, 2005), it is possible that low reading comprehension scores are related to a lack of reading fluency, which is not assessed in PISA. One way of addressing LM/low-SES students' reading difficulties is to attempt to improve their reading fluency and to promote the transition from reading isolated words to reading connected texts. With this study, we propose an attempt to improve reading fluency in LM/low-SES sixth-grade students.

A fluent reader is a reader who reads well, fast enough and with the right expression, which allows an understanding of what is being read. In 1974, LaBerge and Samuels underlined the importance of automaticity in reading performance. They noted that reading presents two levels of treatment: accuracy, where attention is assumed necessary, and automaticity, where it is not. Automaticity allows attention to be devoted to comprehension of the text. Share (1995), through the self-teaching hypothesis, made

3

an interesting proposal on the process of moving from the identification of words to their rapid and autonomous recognition: a series of successful identifications of a word, based on phonological conversions, enables the learner to subsequently recognize it on the basis of stored orthographic information. The combination of accuracy and speed provides the concept of reading rate, measured by the number of words correctly read per minute (NWCR). From the theory of automaticity, we can expect that increases in reading rate should result in increases in comprehension (Fuchs et al., 2001). However, this relation is not always linear. O'Connor (2018) showed in second- and fourth-grade children that a low reading rate is actually correlated with weak comprehension and that the relation between reading rate and comprehension is linear up to a certain level of NWCR, beyond which reading faster no longer contributes to comprehension. This means that reading at an excessively fast rate is less supportive or even counterproductive for comprehension. Moreover, the minimum reading rate for understanding is lower for children with reading disabilities (O'Connor, 2018) or for LM children (Dowd & Bartlett, 2019), suggesting that they read at a rate that enables their comprehension. In contrast, a significant proportion of highly fluent readers in grades 2-10 experience difficulties with higher-order comprehension questions at their grade level (Applegate et al., 2009). This means that the freed-up resources that result from automaticity are not necessarily allocated to comprehension. Reading rate (accuracy and speed) is necessary but not sufficient to develop reading comprehension. Recently, reading prosody has been added to accuracy and speed in the conceptualization of the construct of reading fluency (Kuhn et al., 2010; Rasinski et al., 2009; Schwanenflugel et al., 2004). Reading prosody is the component of reading fluency that makes reading sound like speech. It is reflected by appropriate phrasing, pauses, stress and intonation. This involves raising or lowering the tone at the end of a sentence according to the type of sentence, putting stress on certain words or syllables, making breaks in appropriate places, grouping words in syntactically and semantically coherent chunks and assigning the correct expression based on the meaning of the text and the intention of the author (Dowhower, 1987). Studies suggest that reading prosody and reading comprehension are linked, but the directionality of this relation is debated. For some, reading prosody facilitates reading comprehension (Koriat et al., 2002; Kuhn et al., 2010; Rasinski et al., 2009). Through phrasing and intonational units, reading prosody

4

provides breakpoints in continuous speech that help to ensure that the reader maintains structural units in working memory during a semantic analysis of sentences and allows for the maintenance of meaning (Schwanenflugel et al., 2004). Others claim that prosody is the expression of reading comprehension (Groen et al., 2019) or that the relation is bidirectional (Klauda & Guthrie, 2008; Veenendaal et al., 2016). Furthermore, the nature of the components involved in understanding seems to vary according to the level of reading development. The contribution of automaticity in comprehension is prominent in the first grades of school, while the relation between prosody and comprehension becomes stronger in upper elementary and secondary grades (Fuchs et al., 2001; Groen et al., 2019; Kuhn & Stahl, 2003; Rasinski et al., 2009; Schwanenflugel et al., 2004; Yildirim et al., 2019). In this paper, we refer to the traditional conception of reading fluency measured by the NWCR as reading rate, and we use the term reading fluency for the more current conception that incorporates prosody (see Figure 1).
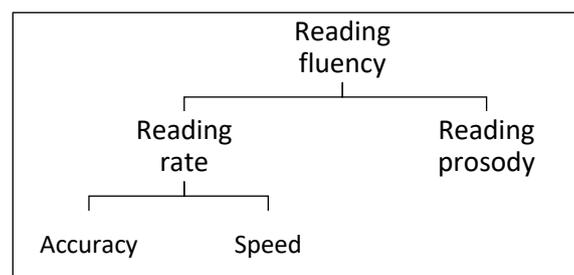

Fig. 1 Components of reading fluency used in this paper

We found few studies that investigated the development of reading fluency in upper elementary LM students. Crosson and Lesaux (2010) showed that in fifth-grade students, rapid word reading and good decoding skills were associated with faster text reading. Furthermore, as with L1 students, NWCR was related to reading comprehension in third- and fifth-grade LM students (Wiley & Deno, 2005). However, this relationship seemed to be more present in students with well-developed listening comprehension skills (Crosson & Lesaux, 2010; Riedel, 2007). Other studies have looked for predictors of reading comprehension in LM students. Gough & Tunmer's (1986) Simple View of Reading is relevant to LM students' reading comprehension: at the beginning of reading comprehension development, word decoding has a prominent role and allows resources to be freed up for comprehension. As students

5

become more competent, the role of word decoding decreases, and listening comprehension plays a greater role (Verhoeven & van Leeuwe, 2012). Although effective decoding is a necessary condition for reading comprehension (Farnia & Geva, 2013; Verhoeven & van Leeuwe, 2012), listening skills determine reading comprehension through the understanding of the meaning of words and sentences, the identification of text structures and the making of inferences. Students with better knowledge of the language of instruction develop better comprehension in upper grades due to better vocabulary knowledge (Droop & Verhoeven, 2003; Lesaux & Kieffer, 2010) or syntactic skills (Farnia & Geva, 2013; Geva & Farnia, 2012).

Fluency intervention studies for struggling readers mainly focus on young children and L1 students. Fluency intervention studies among struggling students in the upper grades are rare, especially for LM or low-income background students. Wanzek & Roberts (2012) studied the effect of multicomponent interventions, including instruction in word recognition, vocabulary and comprehension with an emphasis on either word recognition or explicit comprehension, with struggling fourth-grade readers from economically disadvantaged and linguistically diverse backgrounds. No significant differences were found in word reading, word attack, vocabulary and comprehension measures between the intervention groups or a control group receiving school-implemented interventions. However, students with English-limited proficiency performed better on word attack and word recognition than other students. Lovett et al. (2008) proposed intervention programs focused on teaching basic word identification and decoding skills to struggling readers from grades 2-8. Compared to a control program, students who received intervention showed significant gains in reading achievement. This study took place in an urban school with English-as-a-first-language children and English-as-a-second-language children. Students showed equivalent results at posttest regardless of language status. Denton et al. (2008) focused on LM and/or economically disadvantaged students in grades 6-8 who had severe deficits in reading and compared students engaged in a small-group multicomponent reading intervention to students who received the school's typical instruction. Neither group demonstrated significant progress in word reading, reading rate and reading comprehension. These authors highlighted the difficulties in

finding effective programs to improve reading achievement in older LM/low-SES students. We found no studies that investigated the effects of a reading intervention on prosody in LM/low-SES students.

The present study

Little is known about what interventions are most effective for older LM/low-SES students with reading difficulties. Considering the knowledge on reading fluency and the gains that can be obtained with reading interventions in L1 students (Wanzek et al., 2010), we implemented an experimental design integrating an intervention program focused on reading fluency. We chose to address sixth-grade LM/low-SES students because we felt it was important to look for a way to improve students' reading fluency when the curricula demands for processing texts are higher and a few months before they enter high school. The purpose of this study was to investigate the effectiveness of a reading-aloud training program (RATP) in improving reading rate and reading prosody in a group of low-achievement sixth-grade LM/low-SES readers. In addition, we assessed the impact on reading comprehension.

The RATP is a multicomponent approach developed specifically for this research based on best practices reported in the literature: **repeated reading (RR)**, **modeling prosody**, **assisted wide reading**, and **guided comprehension**. Literature reviews indicate that RR facilitates accuracy, reading rate and comprehension of practiced passages for students across ages and disability status (Chard et al., 2002; Kuhn & Stahl, 2003; Lee & Yoon, 2017; Stevens et al., 2017; Therrien, 2004; Wexler et al., 2008). In the RATP, **RR** instruction requires students to read the same passage until an established reading rate criterion is met, with explicit corrective feedback from a tutor; it is acknowledged that this type of feedback improves reading rate (Therrien, 2004; Wexler et al., 2008). According to Kuhn (2004), readers are able to employ prosodic elements only as they become aware of the connection between written and oral language. Among the components of prosody, because phrasing is closely linked to the grammatical structure of the text, is quite easy to acquire, as is intonation, which can be cued by punctuation. Expressivity, which implies a deep understanding of the text, is most difficult to learn (Godde et al., 2020). However, for many children, the transfer between syntactic or prosodic knowledge

7

from oral language to written language is not automatic. Strategies should be specifically taught for organizing written text into syntactic structures, intonation contours and stress, i.e., for achieving the appropriate prosody (Kuhn et al., 2010; LeVasseur et al., 2008; Rasinski et al., 2009; Schwanenflugel et al., 2004). In the RATP, the tutor helps the student produce appropriate prosody by **modeling prosody** and by providing performance feedback on the components of prosody. Although studies have shown that RR can lead to improvement in students' reading prosody in addition to word-level automaticity (Dowhower, 1987; Kuhn, 2004; Zimmerman et al., 2019), other studies suggest that prosodic skills must be explicitly taught because training in automaticity does not necessarily improve these skills (Ardoin et al., 2013; Calet et al., 2017). The RATP also includes **assisted wide reading**, another practice used to improve reading fluency (Ardoin et al., 2016; Wexler et al., 2008). Wide reading has the advantage of motivating students who follow a story and its protagonists, expands the breadth of vocabulary encountered and provides the opportunity for students to read the same words in multiple contexts. Finally, the RATP uses a procedure of **guided comprehension**, which provides explicit and systematic explanation of unknown vocabulary and the inclusion of questions to engage children in deep processing of the meaning of the text (comprehension questions). Assisted wide reading combined with explanations of vocabulary and comprehension questions could counteract the risk of repeated reading being associated with reading too fast for comprehension (Applegate et al., 2009).

Methods

Participants

The study was conducted in a school located in the French-speaking part of Belgium that had high linguistic diversity as well as the lowest socioeconomic index (ISE). In this part of Belgium, each school is assigned an ISE based on variables related to income per capita, diploma level, unemployment rate and professional activities in the students' household (Fédération Wallonie-Bruxelles, 2020). This mechanism is an attempt to regulate school and social disparities by differentiating the allocation of school funding.

Screening procedure

Two years of recruitment were needed to establish the final sample of participants. In the first year, 43 sixth-grade students from several classes from the same school were screened using a reading rate task (see "Materials and Procedures"). We used this task because it is generally acknowledged that reading rate is one of the fundamental criteria that distinguishes proficient readers from struggling readers (Jenkins et al., 2003). From these 43 students, the 24 students who presented the lowest scores were selected to complete the individual assessments. In the second year of recruitment, 42 students followed the same procedure, and 30 students were selected. Thus, in total, 54 students formed the final sample. The criterion that determined the number of selected students came from an organizational constraint and corresponded to the number of tutors available for the RATP.

Participants and their parents gave their written informed consent prior to their inclusion in the study. The study was conducted according to the ethical guidelines of the ethics committee of the Faculty of Psychology, Speech Therapy and Education Sciences of the University of Liège.

Sociodemographic characteristics

The students were 11 years old (M = 11.11 years; SD = 7.26 months) and included 28 girls and 26 boys from linguistically diverse backgrounds. Students reported speaking at home only French (27.8%), French and one or several other languages (53.7%) or only a language other than French (18.5%). The mother tongue of the students' mothers was French for 26% of students, Arabic for 33.4% and Russian for 13%. The remaining 27.6% of reported mother tongues were Armenian, Romanian, Bengali, Italian, Kurdish, Dutch, Rif, Rwandan, Susu, Spanish, Chechen, and Albanian.

With regard to socioeconomic level, 35.2% of the fathers and 72.2% of the mothers were unemployed, while 35.2% of the fathers and 18.5% of the mothers had low-paying jobs. The majority of parents had a low level of education: 37% of the fathers and 36.2% of the mothers reported having an elementary school level, and 24.1% of the fathers and 31.5% of the mothers reported having a lower secondary school education (in Belgium, this is equivalent to the first three years after elementary school).

9

Materials and Procedures

The 54 students selected for the reading rate task were assessed with different tasks measuring reading rate, reading prosody, vocabulary, and reading comprehension. An arithmetic control task was administered to ensure that outcomes were intervention specific. Students were pretested in February and posttested in May at the end of the RATP. Tasks were the same at pre- and posttests. Screening and pretest outcomes were used as a baseline and enabled matching between the intervention and control groups. Age, vocabulary, nonverbal intelligence and parents' employment and level of education completed the matching between the two groups of 27 students each. The conditions (intervention and control) were randomly assigned to each group. All students completed the study.

Measures

Reading

Reading fluency measures:

    a)   Reading rate

Standard reading task. The *Alouette* test (Lefavrais, 2005) is a standard, commonly used task in which the NWCR (number of words correctly read/reading time x 60) is calculated. We asked the students to read aloud a 265-word text as rapidly and as accurately as possible with a 3-minute time limit. The text is meaningless and prevents readers from using contextual information to compensate for their reading difficulties. As a result, this test assesses pure decoding.

Reading rate task (Jacquier-Roux et al., 2010). We asked the students to read aloud a text as rapidly and as accurately as possible. We stopped the reading after one minute. The score was the NWCR (number of words read in one minute minus the number of errors). To avoid a training effect, we used two different texts: "Le géant égoïste" was used for screening and as a pretest measure, and "Monsieur Petit"

was used in the posttest. The second text was easier than the first according to established norms ("Le géant égoïste": M = 135, SD = 35; "Monsieur Petit": M = 141, SD = 32).

b) Reading prosody

Induced prosody. To assess reading prosody induced by an instruction, we used the Multidimensional Fluency Scale (MFS) (Rasinski, 2004). The MFS is divided into four subscales, *Expression and Volume, Phrasing, Smoothness,* and *Pace,* each of which is scored between 1 (beginner) and 4 (master) to calculate a total score (for details, see appendix 1). We asked the students to accurately read aloud a text, paying attention to punctuation and providing an expressive reading as if they were reading to an audience. We stopped them after 1 minute. A narrative text ("Le roi est nu") was specifically chosen. The text consisted of 198 words and included declarative and exclamative sentences and varied punctuation. Individual students' readings were recorded. Two different assessors who had previously agreed on how to score listened to each student's oral reading and assessed the prosodic quality of the reading using the MFS. Given the high correlation between the results obtained by the two evaluators at pretest (Pearson = 0.68; BF10 = $1.10e^6$), we could consider this prosody assessment procedure to be valid. We performed statistical analyses on the results obtained by one of the two evaluators. In addition to the administration of the MFS scale, we calculated the NWCR obtained in this task to observe this variable in an ecologic task, i.e., without asking the students to read quickly.

Spontaneous prosody. We created this measure a posteriori. We used the records obtained with the reading rate task, and we applied the MFS scale. Because students sometimes adopt an overly fast rate in reading rate tasks, *Pace* was not included in the total score. In the absence of instruction on prosody and given the instruction that encouraged speed on the reading rate task, we wanted to assess whether, after the intervention, trained students spontaneously provided more adequate reading prosody.

Reading comprehension:

11

Sentence reading comprehension. "L3" (Lobrot, 1980) is a silent reading task in which students complete 36 unfinished written sentences by selecting one word from five written proposals, completing as many of the sentences as possible within a 5-minute period.

Vocabulary

Receptive vocabulary. The French version of the Peabody Picture Vocabulary Test (Dunn et al., 1993) was used. The students were asked to select the picture corresponding to a word out of 4 options. Since the students in the study were assumed to present a low lexical level, the starting point of the task was set one year lower than the actual age of the students. Form A was administered at pretest and form B at posttest.

Productive vocabulary. The Expressive One-Word Picture Vocabulary Test (Martin & Brownell, 2000), translated into French, is a picture-naming test. Item 146 ("medical prescription") was removed from the task because the corresponding picture is not used in Belgium. Task administration stopped after six consecutive incorrect responses.

Intervention vocabulary. The intervention, described below, was based on reading a book. During reading, certain words of rarer frequency were systematically explained to the students so that they could acquire the meaning of these words. To assess the effectiveness of this procedure, we compared the knowledge of a sample of 34 of these words (taken only from the chapters all the students read) (list A) to the knowledge of 34 words not encountered in the book and matched on frequency and length (list B). We asked the students to choose from four options the one corresponding to the definition read by the experimenter. First, for the first cohort, we administered the task only to the intervention group. The comparison between the knowledge of encountered words (list A) and nonencountered words (list B) made it possible to evaluate the effect of the vocabulary intervention in this group. For the second cohort, we decided to administer the task to both the control and intervention groups. We expected the intervention group to progress on the list of encountered words, whereas the control group would not progress on either list.

12

Nonverbal intelligence

WISC-IV matrices (Wechsler, 2005). We used the standardized procedure for the administration and score calculation.

Control task

Arithmetic. The Tempo Test Rekenen Test (De Vos, 2004) is an arithmetic fluency task that requires the student to solve as many mathematical operations as possible in one minute by operation (addition, subtraction, multiplication and division). The score is the number of correct responses for the four operations.

Test administration procedures

A project coordinator administered the reading rate task. Students from the psychology or language therapy department of the University of Liège called the tutors, administered the other tasks and delivered the RATP. For task administration, each tutor was assigned two students, one from the intervention group and one from the control group, who differed at pretest and posttest. The tutors did not evaluate the student they accompanied during the intervention program, so they were blinded to the student's condition. Prior to assessing students, tutors were trained to test. Oral measures were recorded.

Intervention Implementation

The RATP was a one-to-one intervention implemented 3 days per week in 30- to 40-min sessions for a period of 7 weeks between February and May. Tutors were assigned a student and followed a well-defined intervention procedure. They provided intervention during the regular school day in a setting outside of the classroom. The project coordinator accompanied the tutors throughout the entire procedure.

The RATP: a multicomponent intervention

To make the program attractive and motivating, we worked with a reading book that we felt could be enjoyable for the students: a nice and informative story whose level of difficulty was appropriate and illustrated with pictures. The book "Picouic et Tigrelin en Europe" (Brehon & Nemo, 2011) tells the story of two friends, a dog and a cat, who discover Europe while traveling. It is divided into 62 chapters.

One session of the RATP is composed of two parts: "Assisted wide reading and guided comprehension" (focused on accuracy of reading and comprehension) and "Repeated reading and modeling prosody" (focused on reading rate and prosody).

"Assisted wide reading and guided comprehension": This part of the RATP took 20 to 25 minutes. At the beginning of each session, the tutor simply asked the student to read the text aloud. When there was a decoding error, if the student did not self-correct, the tutor corrected it, and then the student had to reread the whole sentence containing the word. Emphasis was also on comprehension; when the student read a word written in italics (more or less 10 by chapter), the tutor defined it according to the standard definition given at the end of the chapter. If the student seemed to not understand certain words, the tutor explained them. When the student finished reading one chapter, the tutor asked him or her open-ended comprehension questions and proposed cloze sentences about the text. If the student was unable to answer, he or she could consult the text and, if necessary, the tutor told him or her in which part of the text to find the answer. According to their reading speed, the students read one or two chapters of the book per session, which were one or two pages long. At the next session, the student picked up where he or she previously stopped.

"Repeated reading and modeling prosody": This part took 10 to 15 minutes. We used the same book as the one previously read during the assisted wide reading. Therefore, the students already had the opportunity to become familiar with the words and the story. We divided the chapters of the book into passages of approximately one hundred words each. Starting with the beginning of the book, the student read a passage and followed a sequence of procedures until established criteria were met. First, a) the student read the passage orally for the first time while being timed by the tutor, with the following instruction: read accurately, quickly and with respect for punctuation. Reading errors were pointed out

14

by the tutor, and the student had to reread. The student constructed accurate reading practices by rereading the passage until he or she reached a predetermined reading rate goal, which was established at the beginning of the intervention based on the score obtained on the reading rate task and was increased by approximately 10% for each passage read correctly. Thus, a fixed performance criterion expressed as a rate was chosen according to the contributions of the meta-analysis of Therrien (2004) that indicated that a fixed performance criterion produced a greater effect than a fixed number of rereads. Then, b) the student read with the appropriate prosody. For this purpose, the tutor read aloud the previously trained passage by modeling a quality reading of the text with exemplary phrasing, expression and pace and told the student to read as the tutor had just read, with the same expression, respect for punctuation and pauses. During the first sessions and as long as necessary, the tutor explicitly explained or recalled the links between punctuation and intonation and phrasing in clause and sentence units or emphasizing a single word to help convey meaning. To proceed to the next passage, the student had to meet three criteria: achieve the reading goal rate, have no reading errors, and read with appropriate prosody. At the next session, the student picked up where he or she previously stopped.

Control group

While the intervention group was provided with an intervention outside of the classroom during instructional time, the control group received traditional reading instruction with the teacher. It was essentially silent reading with comprehension questions.

Statistical analysis

The statistical analysis conducted in the present study adopted a Bayesian and model comparison approach. All the analyses were conducted with version 0.9.2 of the JASP software package, using default settings for Cauchy prior distribution. To interpret the Bayes factors, indicative guidelines proposed by Jeffreys (1961) were used: $BF < 1$ = no evidence, $1 < BF < 3$ = anecdotal evidence, $3 < BF < 10$ = moderate evidence, $10 < BF < 30$ = strong evidence, $30 < BF < 100$ = very strong evidence and $BF > 100$ = extreme/decisive evidence for the presence (or absence) of a given effect.

15

When reporting BFs, BF10 indicates evidence in favor of a specific variable/model against the null model, and BF01 indicates the reverse evidence.

Results

Reading rate screening task

The mean score of the 54 selected participants was low (M = 95.72; SD = 18.50), corresponding to a standard score between percentiles 10 and 15, while the maximum score (131.00) corresponded to percentile 50 and the minimum score (55.00) was below percentile 5.

Pretest results and group matching

Pretest performance scores and sociodemographic information for the intervention and control groups are presented in Tables 1 and 5. Bayesian independent samples t test and chi^2 analyses did not reveal evidence of differences between the two groups on any measure except for fathers' employment. This difference was not limiting for our experimental paradigm because it was not in favor of the intervention group and could not explain a possible performance advantage for this group.

Table 1 Parents' level of education and parents' employment - Descriptive statistics for intervention and control groups and comparisons

	Intervention (%)	Control (%)	Intervention (%)	Control (%)
Level of education	Mother's		Father's[1]	
1	25.9	44.4	40.7	33.3
2	40.7	22.2	18.5	29.6
3	14.8	18.5	11.1	14.8
4	11.1	7.4	14.8	0.0
5	7.4	3.7	11.1	14.8
Test chi^2 (BF$_{10}$)	0.26		0.36	
Employment	Mother's[3]		Father's[4]	
1	70.4	74.1	33.3	37.0
2	18.5	3.7	11.1	3.7
3	11.1	3.7	44.4	11.1
4	0.0	7.4	0.0	7.4
5	0.0	3.7	3.7	11.1
6	0.0	3.7	0.0	11.1
7	0.0	0.0	3.7	7.4
Test chi^2 (BF$_{10}$)	1.93		6.91	

Note: Level of education: 1: Elementary school; 2: Lower secondary school; 3: Higher secondary school; 4: High school; 5: University - Employment: 1: Unemployment; 2: Elementary occupations; 3: Craft and related trades workers; 4: Service and sales workers; 5: Clerical support workers; 6: Technicians and associate professionals; 7: Professionals; [1]3.8% (Intervention and 7.5% (Control) of missing values; [3]3.7% (Control) of missing values; [4]3.8% (Intervention) and 11.2% (Control) of missing values

Comparison of pre- and posttest scores

To examine the effect of the intervention, we conducted Bayesian repeated-measures ANOVAs on every variable with time (pretest vs. posttest) as the within-participant factor and condition (intervention vs. control) as the between-participant factor. Three effects were considered: the condition effect (intervention vs. control), time effect (pre- vs. posttreatment), and condition-by-time interaction effect. The main goal of the intervention was the improvement of reading rate and prosody skills in students who trained with the RATP. Additionally, the impact on reading comprehension was assessed. Given the short duration of the intervention, we did not expect improvement in vocabulary performance except in the intervention vocabulary task. Arithmetic control task performance was not supposed to improve further in the intervention group. The results of the posttest performance are presented in Table 5.

Reading

Reading fluency measures:

a) Reading rate

Standard reading test: Alouette (NWCR): Bayesian analysis revealed decisive evidence for an effect of time ($BF_{10} = 3.76e^5$; $BF_{01} = 2.66e^{-6}$; error% = 1.18) greater than the evidence for the interaction between time and condition ($BF_{10} = 5.65e^4$; $BF_{01} = 1.77e^{-5}$; error% = 3.42) for the NWCR score.

Reading rate task (NWCR): At the pretest, the intervention and control groups presented an average performance corresponding to a standard score between the 10th and 15th percentiles. At the posttest, the intervention group presented an average performance in the 30th percentile, and the control group

17

presented an average performance between the 20th and 25th percentiles. However, since Bayesian analysis revealed decisive evidence for an effect of time ($BF_{10} = 1.32e^{17}$; $BF_{01} = 7.59e^{-18}$; error% = 2.29) greater than the evidence for the interaction between time and condition ($BF_{10} = 5.74e^{16}$; $BF_{01} = 2.60e^{-17}$; error% = 3.67) on the NWCR, we cannot conclude that the intervention group outperformed the control group.

b) Reading prosody

Induced prosody: Bayesian analysis revealed evidence of the effects of time ($BF_{10} = 1.43e^{10}$; $BF_{01} = 6.97e^{-11}$; error% = 1.17) and condition ($BF_{10} = 7.24$; $BF_{01} = 0.14$; error% = 0.77) on performance and decisive evidence of the interaction between condition and time ($BF_{10} = 1.97e^{13}$; $BF_{01} = 5.07e^{-14}$; error% = 3.55) (Table 2). As shown in Figure 2, the intervention group outperformed the control group.

Table 2 Results of the Bayesian repeated-measures analysis of variance on the induced prosody score

Model	P(M)	P(M?data)	BF_M	BF_{10}	Error %
Null model (incl. subject)	0.20	$5.02e^{-14}$	$2.01e^{-13}$	1.00	
COND	0.20	$3.64e^{-13}$	$1.45e^{-12}$	7.24	0.77
TIME	0.20	$7.20e^{-4}$	0.003	$1.43e^{10}$	1.17
TIME + COND	0.20	0.01	0.04	$1.82e^{11}$	2.42
TIME + COND + TIME * COND	0.20	0.99	402.03	$1.97e^{13}$	3.5

Note. TIME = time effect; COND = intervention condition effect; P(M) = prior probability of each model; P(M?data) = model's posterior probability; BF = Bayes factor; BF_M = change from prior to posterior model odds; BF_{10} = Bayes factor of each model in comparison to the null model; error % = estimate of the estimation error of the BFs
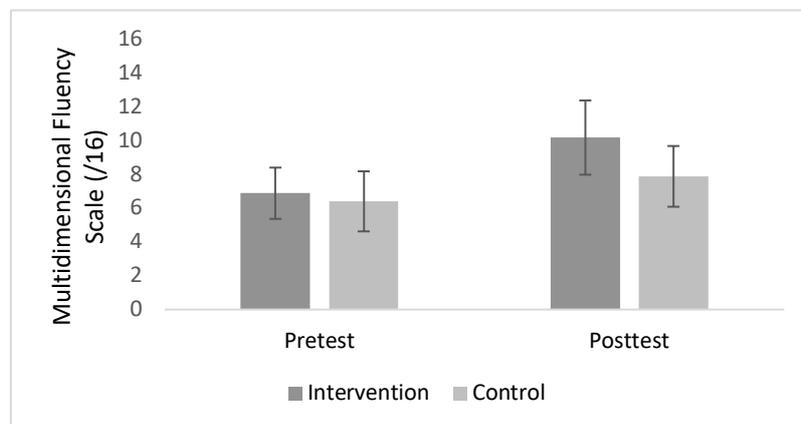

18

Fig. 2 Induced total prosody scores at pretest and posttest by condition (intervention vs. control)

The four subcategories of the MFS indicate decisive evidence for the interaction between time and condition for *Expression and volume* ($BF_{10} = 9.45e^5$; $BF_{01} = 1.06.e^{-6}$; error% = 4.56), *Phrasing* ($BF_{10} = 4.71e^{13}$; $BF_{01} = 2.12e^{-14}$; error% = 2.17), and *Pace* ($BF_{10} = 2.64e^6$; $BF_{01} = 3.79e^{-7}$; error% = 3.01) but not for *Smoothness* ($BF_{10} = 0.54$; $BF_{01} = 1.84$; error% = 1.98) (Table 4).

The additional scoring in this task, i.e., the NWCR, did not show different results for the two groups, with decisive evidence of an effect of time ($BF_{10} = 2.74e^4$; $BF_{01} = 3.65e^{-5}$; error% = 1.19) greater than evidence of the interaction between time and condition ($BF_{10} = 7.58e^3$; $BF_{01} = 1.32e^{-4}$; error% = 3.19).

Spontaneous prosody: Bayesian analysis revealed evidence of the effects of time ($BF_{10} = 2.52e^7$; $BF_{01} = 3.96e^{-8}$; error% = 0.86) and condition ($BF_{10} = 3.94$; $BF_{01} = 0.25$; error% = 1.02) on performance and decisive evidence of the interaction between condition ($BF_{10} = 1.61e^9$; $BF_{01} = 6.20e^{-10}$; error% = 3.05) (Table 3). As shown in Figure 3, the intervention group outperformed the control group.

Table 3 Results of the Bayesian repeated-measures analysis of variance on the spontaneous prosody score

Model	P(M)	P(M?data)	BF_M	BF_{10}	Error %
Null model (incl. subject)	0.20	$5.54e^{-10}$	$2.22e^{-9}$	1.00	
COND	0.20	$2.18e^{-9}$	$8.73e^{-9}$	3.94	1.02
TIME	0.20	0.01	0.06	$2.52e^7$	0.86
TIME + COND	0.20	0.09	0.40	$1.65e^8$	2.35
TIME + COND + TIME * COND	0.20	0.89	33.90	$1.61e^9$	3.05

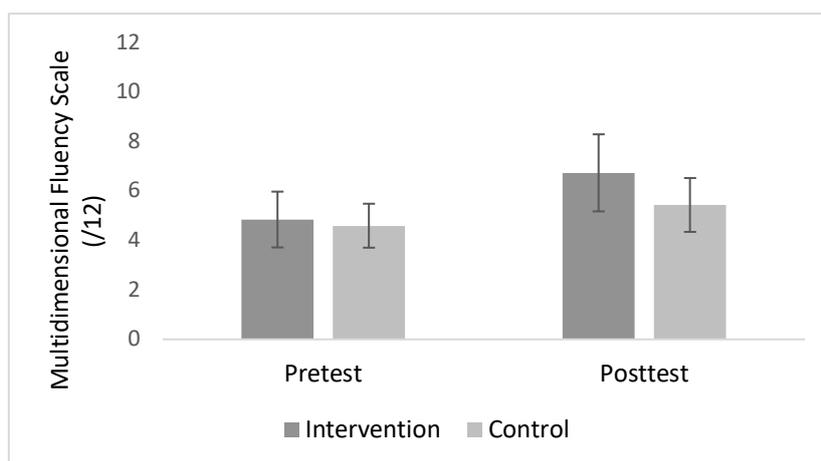*Note.* See Table 2 for species code.

Fig. 3 Spontaneous total prosody scores at pretest and posttest by condition (intervention vs. control)

The three subcategories of the MFS indicate decisive evidence of the interaction between time and condition only for *phrasing* ($BF_{10} = 5.66e^8$; $BF_{01} = 1.76e^{-9}$; error% = 1.74) (Table 4).

Table 4 Descriptive statistics and levels of evidence for ANOVAs of the four subcategories and the total score of the evaluation of induced prosody and spontaneous prosody

	Induced prosody				
	Pretest		Posttest		
	Intervention	Control	Intervention	Control	
	Mean (SD) *min-max*	Mean (SD) *min-max*	Mean (SD) *min-max*	Mean (SD) *min-max*	*LE[1]*
Expression and volume	1.48 (0.64) *1-3*	1.44 (0.51) *1-2*	2.41 (0.69) *1-4*	1.81 (0.83) *1-4*	*decisive*
Phrasing	1.81 (0.56) *1-3*	1.59 (0.57) *1-3*	2.89 (0.70) *2-4*	2.15 (0.46) *1-3*	*decisive*
Smoothness	1.93 (0.55) *1-3*	1.89 (0.58) *1-3*	2.22 (0.64) *1-4*	2.04 (0.34) *1-3*	*no*
Pace	1.67 (0.73) *1-3*	1.48 (0.64) *1-3*	2.66 (0.92) *1-4*	1.89 (0.89) *1-4*	*decisive*
Total	6.89 (1.53) *4-10*	6.41 (1.78) *4-11*	10.19 (2.20) *6-14*	7.89 (1.80) *5-12*	*decisive*
	Spontaneous prosody				
Expression and volume	1.33 (0.48) *1-2*	1.49 (0.36) *1-2*	1.74 (0.94) *1-5*	1.44 (0.58) *1-3*	*no*
Phrasing	1.18 (0.55) *1-3*	1.66 (0.55) *1-3*	2.89 (0.75) *2-5*	2.07 (0.67) *1-3*	*decisive*
Smoothness	1.70 (0.42) *1-3*	1.78 (0.42) *1-2*	2.11 (0.58) *1-5*	1.93 (0.27) *1-5*	*no*
Total	4.85 (1.14) *3-7*	4.59 (0.89) *3-6*	6.74 (1.56) *5-12*	5.44 (1.09) *4-7*	*decisive*

Note. [1]Level of evidence of interaction between time and condition

Reading comprehension:

Sentence reading comprehension: Bayesian analysis indicated decisive evidence of an effect of time ($BF_{10} = 1.84e^8$; $BF_{01} = 5.43e^{-9}$; error% = 1.46) greater than the evidence of the interaction between time and condition ($BF_{10} = 9.81e^7$; $BF_{01} = 1.02e^{-8}$; error% = 3.63).

Vocabulary

Receptive vocabulary: Bayesian analysis did not indicate any time, condition or interaction effects ($BF_{10} = 0.08$; $BF_{01} = 12.15$; error% = 1.07).

Productive vocabulary: Bayesian analysis indicated decisive evidence of an effect of time ($BF_{10} = 9.35e^2$; $BF_{01} = 0.01$; error% = 1.24) greater than the evidence of the interaction between time and condition ($BF_{10} = 6.98e^2$; $BF_{01} = 0.01$; error% = 3.98).

Intervention vocabulary: We first compared the progress of the intervention group on lists A and B. Bayesian paired sample t tests indicated decisive evidence supporting better results in the posttest for list A ($BF_{10} = 14.28e^4$; $BF_{01} = 7.00e^{-6}$; error% = $2.18e^{-9}$) and equivalent results between the pretest and posttest for list B ($BF_{10} = 0.70$; $BF_{01} = 1.42$; error% = $2.24e^{-5}$). We also compared the results of the control and intervention groups. Concerning the control group, ANOVAs were conducted only with the second cohort. For list A, the analysis revealed decisive evidence of the effects of time ($BF_{10} = 2.27e^4$; $BF_{01} = 4.38e^{-5}$; error% = 1.06) and condition ($BF_{10} = 12.36$; $BF_{01} = 0.08$; error% = 0.81) on performance and decisive evidence of the interaction between condition and time ($BF_{10} = 1.01e^7$; $BF_{01} = 9.95e^{-8}$; error% = 2.63). No effect of time ($BF_{10} = 0.56$; $BF_{01} = 1.79$ error% = 1.50) or condition ($BF_{10} = 0.45$; $BF_{01} = 2.22$; error% = 0.81) was found for list B.

Table 5 Pre- and posttest results and sociodemographic information - Descriptive statistics, comparisons and levels of evidence for ANOVAs

	Pretest			Posttest		
	Intervention Mean (SD)	Control Mean (SD)	BF_{10}[4]	Intervention Mean (SD)	Control Mean (SD)	LE[5]

	min-max	min-max		min-max	min-max	
Reading						
Reading fluency measures						
Reading rate						
Alouette (nwcr)	81.18 (18.78)	83.72 (21.48)	*0.30*	93.23 (22.06)	94.79 (23.13)	*no*
	55.33-113.70	*40.33-132.81*		*56.66-146.34*	*58.00-135.70*	
Reading rate (nwcr)[1]	95.41 (18.30)	96.04 (19.04)	*0.28*	129.33 (29.08)	123.81 (25.89)	*no*
	55-131	*55-130*		*68-191*	*75-171*	
Reading prosody						
Induced prosody	6.89 (1.52)	6.41(1.78)	*0.44*	10.19 (2,20)	7.89 (1.80)	*decisive*[8]
(/16)	4-10	4-11		6-14	5-12	
Spontaneous prosody	4.85 (1.13)	4.60 (0.89)	*0.39*	6.74 (1.56)	5.44 (1.09)	*decisive*[8]
(/12)	*3-7*	*3-6*		*5-12*	*4-7*	
Reading comprehension	22.19 (4.61)	22.37 (5.26)	*0.28*	26.56 (6.08)	25.19 (5.86)	*no*
(/36)	*12-31*	*8-30*		*9-36*	*11-34*	
Vocabulary						
Receptive vocabulary[2]	111.93 (20.11)	110.81(26.57)	*0.28*	114.04 (20.10)	107.31 (22.18)	*no*
	72-139	*33-142*		*58-157*	*36-148*	
Productive vocabulary	77.96 (15.17)	73.78 (30.00)	*0.40*	85.37 (17.12)	78.96 (16.46)	*no*
	46-111	*30-94*		*59-122*	*37-110*	
Intervention vocabulary[3]						
List A (%)	44.23 (13.46)	37.84 (11.33)	*0.80*	58.71 (13.01)	40.78 (13.52)	*decisive*
	17.65-79.41	*23.53-58.82*	*2.44*[6]	*29.41-82.35*	*20.59-67.65*	*decisive*[8]
List B (%)	48.91 (12.95)	45.49 (13.47)	*0.40*	45.53 (11.66)	44.90 (12.50)	*no*
	26.47-79.41	*26.47-76.47*		*17.65-64.71*	*14.71-61.76*	*no*[9]
Control task						
Arithmetic	81.19 (15.65)	74.26 (16.09)	*0.79*	83.48 (15.84)	76.78 (16.61)	*no*
	52-125	*50-110*		*55-122*	*46-115*	
Nonverbal intelligence	8.74 (2.33)	8.00 (2.53)				
	4-14	*2-13*				
Sociodemographic information						
Age in years	11.75 (0.54)	12.10 (0.62)	*2.07*			
Sex (% boys)	51.90	48.14	*0.38*[7]			
Mother's mother tongue French (%)	22.20	29.60	*0.44*[7]			

Note. [1]Texts at pretest and posttest differ; [2]EVIP form A at pretest and form B at posttest; [3]Results from the 2 cohorts for intervention group and from the second cohort for control group; [4]Bayesian independent samples t test for mean differences at pretest between intervention and control groups; [5]Level of evidence of interaction between time and condition; [6]Bayesian paired samples t test for mean differences at pretest between lists A and B for intervention group; [7]Bayesian contingency tables tests; [8]Detailed results in Tables 2 and 3; [9]Level of evidence of Bayesian paired samples t test

<u>Reading and linguistic background – LM or French speakers</u>

To determine whether the intervention benefits LM speakers as much as native French speakers, we conducted intervention group repeated-measures ANOVAs with time (pretest vs. posttest) as the within-participant factor and linguistic background (LM vs. French speakers) as the between-participant factor on each of the reading measures. We first conducted Bayesian independent samples t tests on the pretest

results, and these tests did not reveal evidence of differences between the two groups. ANOVAs indicated an effect of time but no effect of linguistic background. We applied the same procedure for the intervention vocabulary task with the same observations (Table 6).

Table 6 Descriptive statistics, comparisons and levels of evidence for ANOVAs according to the linguistic background

	Pretest			Posttest		
	LM (19) Mean (SD) *min-max*	French-sp. (8) Mean (SD) *min-max*	BF_{10}[1]	LM (19) Mean (SD) *min-max*	French-sp. (8) Mean (SD) *min-max*	LE[2]
Reading rate						
The Alouette (nwcr)	81.94 (20.07) *53.33-113.70*	79.37 (16.42) *56.00-101.20*	*0.40*	94.02 (25.29) *56.66-146.30*	91.36 (15.52) *69.00-108.00*	*no*
Reading rate (nwcr)	95.84 (21.06) *55-131*	94.38 (10.10) *79-106*	*0.39*	132.2 (33.66) *68-191*	122.6 (12.36) *103-139*	*no*
Reading prosody						
Induced prosody (/16)	6.89 (1.56) *4-10*	6.87 (1.55) *5-9*	*0.38*	10.11 (2.33) *6-14*	10.38 (2.00) *8-13*	*no*
Spontaneous prosody (/12)	4.79 (1.08) *3-7*	5.00 (1.31) *3-7*	*0.41*	6.58 (1.64) *5-12*	7.12 (1.37) *5-9*	*no*
Reading comprehension (/36)	22.00 (5.27) *12-31*	22.63 (2.67) *18-26*	*0.40*	25.95 (6.56) *9-36*	28.00 (4.84) *22-36*	*no*
Intervention vocabulary						
List A (%)	42.11 (14.74) *17.65-79.41*	49.26 (8.57) *29.41-55.88*	*0.68*	57.89 (14.04) *29.41-82.35*	60.66 (10.77) *44.12-79.41*	*no*
List B (%)	48.45 (14.55) *26.47-79.41*	50.00 (8.75) *32.35-55.88*	*0.39*	44.43 (12.55) *17.65-64.71*	48.16 (9.43) *32.35-64.71*	*no*

Note. [1]Bayesian independent samples t test for mean differences at pretest; [2]Level of evidence of interaction between time and linguistic background

Control task

Arithmetic: Bayesian analysis indicated evidence of an interaction between time and condition ($BF_{10} = 4.53$; $BF_{01} = 0.22$; error% = 3.15) that was not high enough to surpass that of time ($BF_{10} = 3.95$; $BF_{01} = 0.25$; error% = 1.10).

Discussion

The present study was conducted to investigate whether a reading aloud training program (RATP) could improve reading fluency skills (rate and prosody) in a group of low-achievement sixth-grade LM/low-SES students. The RATP was a multicomponent intervention combining repeated reading, assisted wide

23

reading with guided comprehension and modeling prosody. Students selected to participate in the RATP had an average score of 95.72 NWCR on a reading rate-screening task, which corresponds to a standard score between percentiles 10 and 15. There was large variation among them, with scores ranging from 55 to 131. We compared the results from a group of students who received the intervention and a control group matched on different reading and control measures. Specifically, we examined whether the intervention group obtained better results in reading rate (NWCR), reading prosody and reading comprehension.

Despite the intensive intervention program proposed to the intervention group, the intervention and control groups progressed in the reading rate and standard reading tasks without showing a significant difference, and the intervention group did not outperform the control group. We expected that the specific activities implemented in the RATP would produce a better reading rate. The findings in the literature reported that RR as applied in the study, i.e., giving performance feedback on word reading accuracy and reading speed, improves reading speed and the accuracy of practiced and transfer passages (Chard et al., 2002; Kuhn & Stahl, 2003; Lee & Yoon, 2017; Stevens et al., 2017; Therrien, 2004; Wexler et al., 2008). These results might indicate that the intervention had no greater impact on the NWCR than the usual reading activities performed in class (Kuhn & Stahl, 2003), i.e., silent reading with comprehension questions.

Concerning reading prosody, the intervention group, which was trained in prosody with tutor modeling and feedback, showed better results than the control group in the *induced prosody task,* both in the total score and in the subscales of *Expression and volume, Phrasing* and *Pace* but not in *Smoothness*. This means that LM/low-SES students were able to transfer their new prosody skills to a text they had not previously worked on and for which they did not have a model for prosody. In the *spontaneous prosody task,* the intervention group scored better on the *Phrasing* subscale and the total score. This indicates that even if students were not explicitly asked to read with prosody, the intervention group, unlike the control group, brought more appropriate phrasing to the reading, whereas usually when students read faster, they do not plan for breathing or grammatical pauses (Godde et al., 2022). Unlike in the induced

prosody task, they were nevertheless not better in *Expression* because expressivity tends to reduce reading rate. Therefore, students may be inexpressive in an attempt to go faster. The results obtained in the two prosody tasks show that students trained in reading prosody improve their phrasing and expressivity. We can assume that they will do so for any text and that for them, reading prosody and, more particularly, phrasing is automated. This is particularly interesting given the relation between phrasing (Arcand et al., 2014; Veenendaal et al., 2014), expression (Klauda & Guthrie, 2008) and reading comprehension.

The results obtained on the four subscales of the MFC provide another perspective to those obtained in the reading rate task. It is possible that the intervention and control groups arrived at the same result in terms of NWCR by using different processes. The intervention group students did not surpass the control group students perhaps because they developed more adequate prosody. In the reading rate task, in which we asked students to read quickly and well, the intervention group favored prosodic quality over speed; in contrast the control group attached great importance to the criterion of speed. Ardoin et al.'s (2013) study reported quite similar findings. To study the impact of RR on prosody and the impact of the type of feedback given to students on the components of reading fluency, they assigned two different conditions to students. In a prosody condition, they provided instruction and performance feedback regarding the importance of pause lengths following commas and periods; the effect was that students increased their pause lengths. In the other condition, they gave students feedback on their reading speed and accuracy, and students decreased the length of their pauses at these same points. Despite the same number of reading trials during RR and the same drill error correction procedures, the students in the prosody condition did not increase their NWCR on a reading rate task, unlike those in the speed condition.

Additional evidence that reading rate is processed differently in the two groups is provided by the *induced prosody task,* where we asked students to read an unfamiliar text with expression and with respect for punctuation. Students who received RATP received better scores in *Phrasing*, *Expression and volume*, and *Pace*. This subscale measured the reading rate (ranging from slow and laborious to

25

consistently conversational) during sections of minimal disruption (i.e., during sections with little or no errors or hesitations in decoding). The results showed that the control group was more likely to be coded as having slow and laborious reading or moderately slow reading, while the intervention group was more likely to be coded as having moderately slow reading or a mixture of slow and fast reading. This is linked to the better phrasing and better pausing placement, which makes the perceived reading rate faster. While the perceived reading rate seems faster, the NWCR does not improve because longer pauses due to better phrasing compensate for the faster reading rate between the pauses.

We also investigated reading comprehension because the literature converges on the assumption that reading fluency and reading comprehension are related for LM students in higher elementary grades (Geva & Farnia, 2012). No difference was found between the intervention and control groups. This may be due to the comprehension task used in this study, which did not account for the complexity of mechanisms involved in text comprehension. Instead, the reading comprehension task highlighted decoding and semantic skills in single sentences and did not offer many possibilities to use phrasing, intonation or expression to increase comprehension. It would have been better to assess comprehension either on the induced prosody text (and to link comprehension and progress in prosody) or with another text read aloud with an explicit instruction for comprehension (questions or a retelling task). This text could have been used to measure spontaneous prosody and avoid the pitfall of the spontaneous prosody task, which is loss of expressivity in favor of speed. Moreover, to assess the reading comprehension of the connected text, it is important to ensure that LM students know the vocabulary and are able to decode virtually every word. Achieving these two prerequisites is particularly challenging for LM students, who exhibit a high degree of variability.

Finally, we did not expect that progress in vocabulary would be visible on standard vocabulary tests given the short period of intervention. Nevertheless, the study showed that when we provided a definition of words in context, students were able to learn this vocabulary. As expected, arithmetic task scores did not improve differently depending on whether students participated in the RATP. We can

argue that the better results on prosody tasks were RATP-related and were not the product of a general improvement in intervention group skills but were specific to the prosody skills.

Finally, we did not find different results depending on whether the students were native French speakers or from another linguistic background, which is in line with the literature (Lesaux & Kieffer, 2010).

Conclusion

All of the data converge toward the assumption that assessing reading fluency using a scale such as the MFS allows an analysis that better captures the complexity of reading fluency in LM/low-SES students than the NWCR score. In line with suggestions by Kuhn et al. (2010), prosody should be measured whenever reading fluency is measured.

The present data provided empirical support for the proposition that reading prosody is a component of reading that can be taught effectively to struggling LM/low-SES students. These students could benefit from a reading fluency intervention to develop important reading skills such as phrasing, expression and pace in educational language, not only on practiced texts but also on new texts, as discussed. In future research, it would be interesting to replicate these findings and to show whether the improvement in reading prosody, as shown in this study, is associated with an improvement in text reading comprehension with more appropriate tasks than the one we used. It would be necessary to investigate how to replace external tutors to make the process ecological to the classroom, i.e., with peer tutoring. In any case, teachers should place more emphasis on teaching phrasing and expressivity in oral reading to all students, even those with reading difficulties.

References

Applegate, M. D., Applegate, A. J., & Modla, V. B. (2009). "She's my best reader; she just can't comprehend": Studying the relationship between fluency and comprehension. *The Reading Teacher*, *62*(6), 512–521. https://doi.org/10.1598/rt.62.6.5

Arcand, M. S., Dion, E., Lemire-Théberge, L., Guay, M. H., Barrette, A., Gagnon, V., Caron, P. O., & Fuchs, D.

(2014). Segmenting texts into meaningful word groups: Beginning readers' prosody and comprehension. *Scientific Studies of Reading*, *18*(3), 208–223. https://doi.org/10.1080/10888438.2013.864658

Ardoin, S. P., Binder, K. S., Foster, T. E., & Zawoyski, A. M. (2016). Repeated versus wide reading: A randomized control design study examining the impact of fluency interventions on underlying reading behavior. *Journal of School Psychology*, *59*, 13–38. https://doi.org/10.1016/j.jsp.2016.09.002

Ardoin, S. P., Morena, L. S., Binder, K. S., & Foster, T. E. (2013). Examining the impact of feedback and repeated readings on oral reading fluency: Let's not forget prosody. *School Psychology Quarterly*, *28*(4), 391–404. https://doi.org/10.1037/spq0000027

Brehon, G., Nemo, J. (2011). *Manuel de lecture suivie: Picouic et Tigrelin en Europe: CE1*. La Librairie des écoles.

Bricteux, S., Crépin, F., Quittre, V., & Lafontaine, D. (2021). Résultats de PISA 2018 en Fédération Wallonie-Bruxelles: Les jeunes de 15 ans et la lecture. *Les Cahiers Des Sciences de l'Education*, *June*.

Buckingham, J., Wheldall, K., & Beaman-Wheldall, R. (2013). Why poor children are more likely to become poor readers: The school years. *Australian Journal of Education*, *57*(3), 190–213. https://doi.org/10.1177/0004944113495500

Calet, N., Gutiérrez-Palma, N., & Defior, S. (2017). Effects of fluency training on reading competence in primary school children: The role of prosody. *Learning and Instruction*, *52*, 59–68. https://doi.org/10.1016/j.learninstruc.2017.04.006

Chard, D. J., Vaughn, S., & Tyler, B.-J. (2002). A synthesis of research on effective interventions for building reading fluency with elementary students with learning disabilities. *Journal of Learning Disabilities*, *35(5)*, 386–406.

Crosson, A. C., & Lesaux, N. K. (2010). Revisiting assumptions about the relationship of fluent reading to comprehension: Spanish-speakers' text-reading fluency in English. *Reading and Writing*, *23*(5), 475–494. https://doi.org/10.1007/s11145-009-9168-8

Denton, C. A., Wexler, J., Vaughn, S., & Bryan, D. (2008). Intervention provided to linguistically diverse middle school students with severe reading gifficulties. *Learning Disabilities Research & Practice*, *23*(2), 79–89. https://doi.org/10.1111/j.1540-5826.2008.00266.x

De Vos, T. (2004). *Tempo-Test-Rekenen: test voor het vaststellen van het rekenvaardigheidsniveau van de elementaire bewerkingen (automatisering) voor het basis- en voortgezet onderwijs*. Amsterdam: Pearson Education

Dowd, A. J., & Bartlett, L. (2019). The need for speed: Interrogating the dominance of oral reading fluency in international reading efforts. *Comparative Education Review*, *63*(2), 189–212. https://doi.org/10.1086/702612

Dowhower, S. L. (1987). Effects of repeated reading on second-drade transitional readers' fluency and comprehension. *Reading Research Quarterly*, *22*(4), 389. https://doi.org/10.2307/747699

Droop, M., & Verhoeven, L. (2003). Language proficiency and reading ability in first- and second- language learners. *Reading Research Quarterly*, *38*, 78–103.

Dunn, L. M., Theriault-Whalen, C. M., & Dunn, L. M. (1993). Échelle de vocabulaire en images peabody. Adaptation française du Peabody Picture Vocabulary test-revised. Belgique: ATM.

Farnia, F., & Geva, E. (2013). Growth and predictors of change in English language learners' reading comprehension. *Journal of Research in Reading*, *36*(4), 389–421. https://doi.org/10.1111/jrir.12003

Fuchs, L. S., Fuchs, D., Hosp, M. K., & Jenkins, J. R. (2001). Oral reading fluency as an indicator of reading competence: A theoretical, empirical, and historical analysis. *The Role of Fluency in Reading Competence, Assessment, and Instruction*, *8438*, 239–256. https://doi.org/10.4324/9781410608246-3

Geva, E., & Farnia, F. (2012). Developmental changes in the nature of language proficiency and reading fluency paint a more complex view of reading comprehension in ELL and EL1. *Reading and Writing*, *25*(8), 1819–1845. https://doi.org/10.1007/s11145-011-9333-8

Godde, E., Bailly, G., & Bosse, M. L. (2022). Pausing and breathing while reading aloud: development from 2nd to 7th grade in French speaking children. *Reading and Writing*, *35*(1), 1–27. https://doi.org/10.1007/s11145-021-10168-z

Godde, E., Line, M., & Gérard, B. (2020). A review of reading prosody acquisition and development. *Reading and Writing*, *33*(2), 399–426. https://doi.org/10.1007/s11145-019-09968-1

Gough, P. B., & Tunmer, W. E. (1986). Decoding, reading, and reading disability. *Remedial and Special Education*, *7*, 6–10.

Groen, M. A., Veenendaal, N. J., & Verhoeven, L. (2019). The role of prosody in reading comprehension:

Evidence from poor comprehenders. *Journal of Research in Reading*, *42*(1), 37–57.

https://doi.org/10.1111/1467-9817.12133

Jacobs, D., Rea, A., Teney, C., Callier, L., & Lothaire, S. (2009). *L'ascenseur social reste en panne. Les performances des élèves issus de l'immigration en Communauté française et en Communauté flamande* (Issue May). http://hdl.handle.net/2013/ULB-DIPOT:oai:dipot.ulb.ac.be:2013/29756

Jacquier-Roux, M., Lequette, C., Pouget, G., Valdois, S., & Zorman, M. (2010). BALE - Batterie Analytique du Langage Ecrit. *UPMF, Grenoble*.

Jeffreys, H. (1961). *Theory of probability* (3rd ed.). Oxford University Press.

Jenkins, J. R., Fuchs, L. S., Van den Broek, P., Espin, C., & Deno, S. L. (2003). Sources of individual differences in reading comprehension and reading fluency. *Journal of Educational Psychology*, *95*(4), 719–729. https://doi.org/10.1037/0022-0663.95.4.719

Kieffer, M. J. (2011). Converging trajectories: Reading growth in language minority learners and their classmates, kindergarten to grade 8. In *American Educational Research Journal* (Vol. 48, Issue 5). https://doi.org/10.3102/0002831211419490

Kieffer, M. J., & Lesaux, N. K. (2012). Effects of academic language instruction on relational and syntactic aspects of morphological awareness for sixth graders from linguistically diverse backgrounds. *Elementary School Journal*, *112*(3), 519–545. https://doi.org/10.1086/663299

Kim, Y. S., Wagner, R. K., & Lopez, D. (2012). Developmental relations between reading fluency and reading comprehension: A longitudinal study from Grade 1 to Grade 2. *Journal of Experimental Child Psychology*, *113*(1), 93–111. https://doi.org/10.1016/j.jecp.2012.03.002

Klauda, S. L., & Guthrie, J. T. (2008). Relationships of three components of reading fluency to reading comprehension. *Journal of Educational Psychology*, *100*(2), 310–321. https://doi.org/10.1037/0022-0663.100.2.310

Koriat, A., Greenberg, S. N., & Kreiner, H. (2002). The extraction of structure during reading: Evidence from reading prosody. *Memory and Cognition*, *30*(2), 270–280. https://doi.org/10.3758/BF03195288

Kuhn, M. (2004). Helping students become accurate, expressive readers: Fluency instruction for small groups. *The Reading Teacher*, *58*(4), 338–344. https://doi.org/10.1598/rt.58.4.3

Kuhn, M. R., Schwanenflugel, P. J., Meisinger, E. B., Levy, B. A., & Rasinski, T. V. (2010). Aligning theory and

assessment of reading fluency: Automaticity, prosody, and definitions of fluency. *Reading Research Quarterly*, *45*(2), 230–251. https://doi.org/10.1598/rrq.45.2.4

Kuhn, M. R., & Stahl, S. A. (2003). Fluency: A review of developmental and remedial practices. In *Journal of Educational Psychology* ((5th ed.), Vol. 95, Issue 1, pp. 3–21). https://doi.org/10.1037/0022-0663.95.1.3

LaBerge, D., & Samuels, S. J. (1974). Towards a theory of automatic information processing in reading. *Cognitive Psychology*, *6*, 293–323.

Lee, J., & Yoon, S. Y. (2017). The effects of repeated reading on reading fluency for students with reading disabilities: A meta-analysis. In *Journal of Learning Disabilities* (Vol. 50, Issue 2, pp. 213–224). https://doi.org/10.1177/0022219415605194

Lefavrais, P. (2005). *Alouette-R: test d'analyse de la vitesse en lecture à partir d'un texte.* Éditions du centre de psychologie appliquée

Lesaux, N. K., & Kieffer, M. J. (2010). Exploring sources of reading comprehension difficulties among language minority learners and their classmates in early adolescence. In *American Educational Research Journal* (Vol. 47, Issue 3). https://doi.org/10.3102/0002831209355469

LeVasseur, V. M., MacAruso, P., & Shankweiler, D. (2008). Promoting gains in reading fluency: A comparison of three approaches. *Reading and Writing*, *21*(3), 205–230. https://doi.org/10.1007/s11145-007-9070-1

Lobrot, M. (1983). *Batterie d'épreuves pour mesurer l'ORthographe et la LECture*. EAP, Issy-les-Moulineaux

Lovett, M. W., De Palma, M., Frijters, J., Steinbach, K., Temple, M., Benson, N., & Lacerenza, L. (2008). Interventions for reading difficulties: A comparison of response to intervention by ELL and EFL struggling readers. *Journal of Learning Disabilities*, *41*(4), 333–352. https://doi.org/10.1177/0022219408317859

Martin, N., & Brownell, R. (2000). *Expressive One-Word Picture Vocabulary Test.* Academic Therapy

O'Connor, R. E. (2018). Reading fluency and students with reading disabilities: How fast is fast enough to promote reading comprehension? *Journal of Learning Disabilities*, *51*(2), 124–136. https://doi.org/10.1177/0022219417691835

Perfetti, C. (2007). Reading ability: Lexical quality to comprehension. *Scientific Studies of Reading*, *11*(4), 357–383. https://doi.org/10.1080/10888430701530730

Rasinski, T., Rikli, A., & Johnston, S. (2009). Reading fluency: More than automaticity? More than a concern for the primary grades? *Literacy Research and Instruction*, *48*(4), 350–361.

https://doi.org/10.1080/19388070802468715

Rasinski, T. V. (2004). Assessing Reading Fluency. *Pacific Resources for Education and Learning*.

Raudszus, H., Segers, E., & Verhoeven, L. (2021). Patterns and predictors of reading comprehension growth in

first and second language readers. *Journal of Research in Reading*, *44*(2), 400–417.

https://doi.org/10.1111/1467-9817.12347

Riedel, B. W. (2007). The relation between DIBELS, reading comprehension, and vocabulary in urban first-grade

students. *Reading Research Quarterly*, *42*(4), 546–567. https://doi.org/10.1598/rrq.42.4.5

Schwanenflugel, P. J., Hamilton, A. M., Kuhn, M. R., Wisenbaker, J. M., & Stahl, S. A. (2004). Becoming a fluent

reader: Reading skill and prosodic features in the oral reading of young readers. *Journal of Educational*

Psychology, *96*(1), 119–129. https://doi.org/10.1037/0022-0663.96.1.119

Share, D. L. (1995). Phonological recoding and self-teaching: sine qua non of reading acquisition. *Cognition*, *55*,

151–218.

Stevens, E. A., Walker, M. A., & Vaughn, S. (2017). The effects of reading fluency interventions on the reading

fluency and reading comprehension performance of elementary students with learning disabilities: A

synthesis of the research from 2001 to 2014. *Journal of Learning Disabilities*, *50*(5), 576–590.

https://doi.org/10.1177/0022219416638028

Therrien, W. J. (2004). Fluency and comprehension gains as a result of repeated reading: A meta-analysis.

Remedial & Special Education, *25*(4), 252–261.

Veenendaal, N. J., Groen, M. A., & Verhoeven, L. (2014). The role of speech prosody and text reading prosody in

children's reading comprehension. *British Journal of Educational Psychology*, *84*(4), 521–536.

https://doi.org/10.1111/bjep.12036

Veenendaal, N. J., Groen, M. A., & Verhoeven, L. (2016). Bidirectional relations between text reading prosody

and reading comprehension in the upper primary school grades: A longitudinal perspective. *Scientific*

Studies of Reading, *20*(3), 189–202. https://doi.org/10.1080/10888438.2015.1128939

Verhoeven, L., & van Leeuwe, J. (2012). The simple view of second language reading throughout the primary

grades. *Reading and Writing*, *25*(8), 1805–1818. https://doi.org/10.1007/s11145-011-9346-3

Wanzek, J., & Roberts, G. (2012). Reading interventions with varying instructional emphases for fourth graders

with reading difficulties. *Learning Disability Quarterly*, *35*(2), 90–101.

https://doi.org/10.1177/0731948711434047

Wanzek, J., Wexler, J., Vaughn, S., & Ciullo, S. (2010). Reading interventions for struggling readers in the upper

elementary grades: A synthesis of 20 years of research. *Reading and Writing*, *23*(8), 889–912.

https://doi.org/10.1007/s11145-009-9179-5

Wechsler, D. (2005). *The Wechsler Intelligence Scale for Children—Fourth Edition*. London: Pearson

Wexler, J., Vaughn, S., Edmonds, M., & Reutebuch, C. K. (2008). A synthesis of fluency interventions for

secondary struggling readers. *Reading and Writing*, *21*(4), 317–347. https://doi.org/10.1007/s11145-007-

9085-7

Wiley, H. I., & Deno, S. L. (2005). Oral reading and maze measures as predictors of success for English learners

on a state standards assessment. *Remedial and Special Education*, *26*(4), 207–214.

https://doi.org/10.1177/07419325050260040301

Yildirim, K., Rasinski, T., & Kaya, D. (2019). Fluency and comprehension of narrative texts in Turkish students in

grades 4 through 8. *Education 3-13*, *47*(3), 348–357. https://doi.org/10.1080/03004279.2018.1449880

Zimmerman, B. S., Rasinski, T. V., Was, C. A., Rawson, K. A., Dunlosky, J., Kruse, S. D., & Nikbakht, E. (2019).

Enhancing outcomes for struggling readers: Empirical analysis of the fluency development lesson.

Reading Psychology, *40*(1), 70–94. https://doi.org/10.1080/02702711.2018.1555365

	1	2	3	4
Expression and volume	Reads with little expression or enthusiasm in voice. Reads words as if simply to get them out. Little sense of trying to make text sound like natural language. Tends to read in a quiet voice.	Some expression. Begins to use voice to make text sound like natural language in some areas of the text but not others. Focus remains largely on saying the words. Still reads in a quiet voice.	Sounds like natural language throughout the better part of the passage. Occasionally slips into expressionless reading. Voice volume is generally appropriate throughout the text.	Reads with good expression and enthusiasm throughout the text. Sounds like natural language. The reader is able to vary expression and volume to match his or her interpretation of the passage.
Phrasing	Monotonic with little sense of phrase boundaries;	Frequent two- and three-word phrases giving the	Mixture of run-ons, mid-sentence pauses for	Generally well phrased, mostly in clause and

33

	frequent word-by-word reading.	impression of choppy reading; improper stress and intonation that fail to mark ends of sentences and clauses.	breath, and possibly some choppiness; reasonable stress/intonation.	sentence units, with adequate attention to expression.
Smoothness	Frequent extended pauses, hesitations, false starts, sound-outs, repetitions, and/or multiple attempts.	Several "rough spots" in text where extended pauses, hesitations, etc., are more frequent and disruptive.	Occasional breaks in smoothness caused by difficulties with specific words and/or structures.	Generally, smooth reading with some breaks, but word and structure difficulties are resolved quickly, usually through self-correction.
Pace (during sections of minimal disruption)	Slow and laborious.	Moderately slow.	Uneven mixture of fast and slow reading.	Consistently conversational.

Appendix 1 Multidimensional Fluency Scale (Rasinski, 2004) – reproduced with the permission of T. Rasinski