Bilingualism: Language and Cognition

cambridge.org/bil

Tracking reading development in an English language university-level bridging program: evidence from eye-movements during passage reading

Research Article

Supplementary material can be found online at https://doi.org/10.1017/S1366728922000542

Cite this article: Schmidtke D, Rahmanian S, Moro AL (2023). Tracking reading development in an English language university-level bridging program: evidence from eye-movements during passage reading. *Bilingualism: Language and Cognition* **26**, 356–370. https://doi.org/10.1017/S1366728922000542

Received: 7 June 2021
Revised: 14 July 2022
Accepted: 19 July 2022
First published online: 15 August 2022

Keywords:
eye-movements; reading development; international students; bridging programs

Address for correspondence:
Daniel Schmidtke
McMaster University
Department of Linguistics and Languages
MELD Office
L.R. Wilson Hall
Hamilton, Ontario
Canada L8N 1E9
Email: schmiddf@mcmaster.ca

Daniel Schmidtke, Sadaf Rahmanian and Anna L. Moro

Department of Linguistics and Languages, MELD Offfice, L.R. WIlson Hall, McMaster University, Hamilton, ON, Canada

Abstract

Increasing numbers of international students enter university education via English language bridging programs. Much research has overlooked the nature of second language reading development during a bridging program, focusing instead on the development of literacy skills of international students who already meet the language requirement for undergraduate admission. We report a longitudinal eye-movement study assessing English passage reading efficiency and comprehension in 405 Chinese-speaking bridging program students. Incoming IELTS reading scores were used as an index of baseline reading ability. Linear mixed-effects regression models fitted to global eye-movement measures and reading comprehension indicated that despite initial between-subjects differences, within-subject change at each ability level progressed at the same rate, following parallel growth trajectories. Therefore, there was significant overall reading progress during the bridging program, but no evidence that the gap between low and high ability readers either closed or widened over time.

Introduction

As English-dominant universities continue to host growing numbers of international students who use English as an additional language (EALs), there has been a well-documented increase in English language bridging programs (e.g., Ling, Wolf, Cho & Wang, 2014; Redden, 2013; Van Viegen & Russell, 2019). The function of these programs is to improve the English language skills of students who fulfill the academic requirements of the host institution, but who do not meet the language proficiency threshold to study at the undergraduate level. There is currently very little available research on the language skill development of EAL students enrolled in university-level English language bridging programs, particularly those that span an academic year (8 months). A citation search on Web of Science showed that there are currently fewer than 10 scientific articles published on the language development of students enrolled in English language bridging programs in higher education. Research specifically on the development of second language (L2) reading across bridging programs is extremely limited. We addressed this issue by collecting longitudinal eye-movement data on passage reading and reading comprehension in three successive cohorts of L1 Chinese students enrolled in a pre-university English bridging program. Using established global eye-movement measures and comprehension accuracy as signatures of passage reading ability, our research objective was to gain a basic understanding of EAL passage reading behavior and its development during intensive language instruction. We examined whether initial differences in reading ability (i) translate into differences in overall reading ability between EAL students, and (ii) lead to different trajectories of reading ability growth.

Estimates of reading speed in EAL students

A large body of L2 research is devoted to the study of L2 reading ability and its predictors, as found in extensive literature reviews (e.g., Bernhardt, 2011; Grabe, 2009; Koda, 2005) and meta-analyses (e.g., Jeon & Yamashita, 2014; Melby-Lervåg & Lervåg, 2014). Research concerning adult L1 Chinese L2 English bilinguals in university contexts focuses on the interrelations between various L2 reading skills (e.g., Xue & Jiang, 2017; Zhang & Koda, 2012) or on the relationship between L1 and L2 reading proficiency (e.g., Jiang, 2011; Xue, 2021). However, much less studied is the extent of reading development of pre-university adult Chinese learners enrolled in English bridging programs. This is an important topic given the growing body of research showing that L2 skills implicated in reading predict academic performance among international students (e.g., Daller & Phelan, 2013; Daller & Xue, 2009; Harrington &

Roche, 2014; Thorpe, Snell, Davey-Evans & Talman, 2017; Trenkic & Warmington, 2019; Yixin & Daller, 2014).

It is not hard to appreciate how the academic success of EALs could hinge on the fast and accurate processing of text. Efficient reading is an important mechanism for absorbing new information, and is especially important in high stakes, time-constrained scenarios, such as exams or reading from lecture slides. Empirical support for the importance of reading fluency in the context of higher education was provided in a study by Trenkic and Warmington (2019), showing that individual differences in literacy skills (higher literacy skills, speed of written word processing, vocabulary, and spelling) collectively accounted for a staggering 51.1% of the variance in EAL students' academic grades, but only 10.7% of the variance in native English-speaking students' grades.

Given the importance of reading skill for EAL students in undergraduate programs, our FIRST CONTRIBUTION is to gain a basic understanding of the development of reading fluency under conditions when pre-undergraduate bridging program EALs are reading for comprehension. Reading fluency, a term used interchangeably with reading efficiency, is typically gauged by measures that assess the speed at which a reader is able to extract information during text reading. Most of the existing research on university-level EAL students has adopted silent reading rate, expressed in WORDS PER MINUTE (wpm), as a measure of reading fluency. From a meta-analysis (Nakanishi, 2015), and an additional literature search, we identified nine studies that examined change in EAL silent reading rate in university-level educational contexts (Al-Homoud & Schmitt, 2009; Beglar, Hunt & Kite, 2012; Beglar & Hunt, 2014; Huffman, 2014; Lao & Krashen, 2000; McLean & Rouault, 2017; Robb & Susser, 1989; Sakurai, 2015; Suk, 2017). Mean reading rates, average change, and corresponding effect sizes, L1, and reading condition for each study are provided in Table 1. The median reading rate for extensive reading is estimated at 97.27 wpm pre-test and 113.86 wpm post-test (the estimates obtained by Lao & Krashen, 2000 are clear outliers). The median reading rate reported across studies for the intensive reading conditions are slower overall: 93.86 wpm pre-test and 94.79 wpm post-test. Taking together the results of the literature review, L2 reading speed AFTER language instruction is expected to be at most 143 wpm slower than the estimated 238 wpm silent reading rate of an average adult native-English speaker (Brysbaert, 2019b). To the best of our knowledge, no study has estimated the reading efficiency of EAL bridging-program students, let alone change in reading efficiency throughout bridging program instruction.

Patterns of L2 reading ability growth

Our SECOND CONTRIBUTION then, is to characterize the trajectory of passage reading efficiency development of bridging program EALs. As shown in Table 1, available estimates of statistically reliable gains in reading rates range from 8 wpm ($d = 0.41$) to 34 wpm ($d = 1.65$) for extensive reading, and from 5 wpm ($g = 0.76$) to 26 wpm ($d = 1.03$) for intensive reading. Effect sizes of small and large magnitudes were observed in programs that were both short (10 weeks) and long (~30 weeks). It is therefore reasonable to hypothesize that gains in reading efficiency ought to be observed in an EAL bridging program for university students with either a short or long duration.

Critically, each of the studies listed in Table 1 examined AVERAGE CHANGE in reading efficiency after English language training. However, an overlooked possibility is that individual differences in reading skill may influence the rate of change in reading efficiency over time. In the present study, we direct our focus on individual variability in reading gains and how this might be affected by the reading skills that are in place at the outset of the bridging program. At a practical level, examining the relationship between reading skill at the outset of English language training and future reading development is important. This information can be used as a diagnostic to help the early identification of students who may struggle to catch up with their peers during the bridging program.

The available literature on patterns of individual differences in L2 reading ability growth largely concerns the development of comprehension and focuses on earlier stages of literacy acquisition, i.e., between kindergarten and grade eight (see an excellent review in Raudszus, Segers & Verhoeven, 2021). For example, Droop and Verhoeven (2003) found that greater initial vocabulary knowledge is linked to larger reading comprehension gains in L2 Dutch children between third and fourth grade. This result was also found by Lervåg and Aukrust (2010) in a sample of L2 Norwegian children between first and third grade. These patterns indicate that a greater L2 knowledge leads to a wider gap between low and high ability readers later in early L2 literacy development (this pattern is typically referred to as the Matthew Effect in studies of L1 literacy acquisition, e.g., Stanovich, Nathan & Vala-Rossi, 1986).

In later school grades, lower L2 ability leads to greater growth. For example, Raudszus et al. (2021) reported that lower decoding ability and vocabulary knowledge predicted larger gains in L2 Dutch readers between fourth and sixth grade. This pattern was also found by Nakamoto, Lindsey and Manis (2007): Spanish L1s who had lower oral language scores made greater L2 English reading comprehension gains. In a large sample of L2 English readers in the US, Kieffer (2011) found that bilinguals who had low English proficiency in kindergarten made greater gains by eighth grade in measures of English literacy, including in reading comprehension. Additionally, Farnia and Geva (2013) showed that L2 English readers with low vocabulary and syntax skills at Grade 1 made larger reading comprehension gains by sixth grade. These patterns imply that, at this age bracket, lower initial L2 proficiency results in larger growth in reading skill, and a narrower gap between readers of varying ability at later stages of development (a pattern referred to as a compensatory model of development in the L1 literature, e.g., Pfost, Hattie, Dörfler & Artelt, 2014).

In summary, initial status in a range of L2 language abilities might differentially impact subsequent growth in L2 reading ability. The patterns from available studies suggest that greater initial L2 proficiency is linked to either (i) larger subsequent gains, or (ii) smaller subsequent gains. Notably, the above-referenced research on L2 reading trajectories is constrained to the developmental trajectories of young L2 populations, involves populations with a diversity of L1s and L2s, and focuses on reading comprehension as a reading outcome. The current longitudinal study departs from prior work by investigating the reading trajectory profiles in a sample of young adult L1 Chinese bridging program students varying in incoming reading ability. We further extend prior work by examining within-participant change in reading comprehension AND eye-movements during passage reading. Before describing our hypotheses, we provide an overview of the relevant eye-movement research in L2 reading.

https://doi.org/10.1017/S1366728922000542 Published online by Cambridge University Press

Table 1. Available estimates of silent reading rate (words per minute) and silent reading rate change in English L2 readers.

Study with ER/IR contrast	Extensive Reading				Control (Intensive Reading)				Study Duration	L1
	pre	post	Δ	*d*	pre	post	Δ	*d*		
Al-Homoud and Schmitt (2009)	60	93.57	33.57*	1.65	61.62	87.75	26.13*	1.03	13 weeks	Arabic
Beglar et al. (2012) group 1	89.71	97.73	8.02*	0.41	87.54	90.51	2.97	0.22	28 weeks	Japanese
Beglar et al. (2012) group 2	94.5	107.34	12.84*	0.65						
Beglar et al. (2012) group 3	103.09	119.93	16.84*	0.67						
Huffman (2014)	110.59	131.33	20.74	0.86	103.76	103.14	−0.62	0.036	15 weeks	Japanese
Lao and Krashen (2000)	235	327	92*	–	89.92	94.79	4.87	–	14 weeks	Mandarin and Cantonese
McLean and Rouault (2017)	99.38	130.34	30.96*	1.32[b]	97.79	103.05	5.26*	0.76[b]	30 weeks	Japanese
Robb and Susser (1989)	79.31	86.55	7.24[a]	–	78.5	76.55	−1.95[a]	–	1 academic year	Japanese
Suk (2016)	133.29	168.42	35.13*	1.02	147.76	163.29	15.53*	0.5	15 weeks	Korean
Study (without ER/IR contrast)										
Beglar & Hunt (2014) group 1	97.27	130.26	32.99	–					28 weeks	Japanese
Beglar & Hunt (2014) group 2	96.9	115.61	18.71	–						
Beglar & Hunt (2014) group 3	93.31	104.56	11.25	–						
Beglar & Hunt (2014) group 4	90.26	94.76	4.5	–						
Beglar & Hunt (2014) group 5	107.9	103.99	−3.91	–						
Sakurai (2015)	101.12	113.86	12.74	0.36					15 weeks	Japanese

*indicates that a statistically significant effect at $p < 0.05$ was observed for the pre-post contrast.
[a]signifies no test statistic for contrast provided.
[b]signifies Hedge's *g* used as effect size estimate.
-indicates that insufficient information was provided to compute an effect size.

Eye-movements and L2 reading

The present study investigates change in EAL reading efficiency using the eye-tracking method, one of the most reliable behavioral paradigms for studies of authentic reading (Liversedge, Blythe & Drieghe, 2012; Rayner, 1998). The eye-tracking technique allows the researcher to record non-invasively, and with high precision, the location and timing of eye-movements, providing valuable information about the moment-to-moment cognitive processes underlying reading. Eye-movement studies of reading are a growing practice in the fields of applied linguistics and L2 reading research (see reviews in Godfroid, 2020; Conklin, Pellicer-Sánchez & Carrol, 2018), covering topics such as passage reading during studying (Dirix, Vander Beken, De Bruyne, Brysbaert & Duyck, 2020), lexical processing (Philipp & Huestegge, 2015; Schmidtke & Moro, 2021; Whitford & Titone, 2012), incidental vocabulary learning (Godfroid, Ahn, Choi, Ballard, Cui, Johnston, Lee, Sarkar & Yoon, 2018; Pellicer-Sánchez, 2016), the effects of cross-linguistic influence (Cop, Dirix, Van Assche, Drieghe & Duyck, 2017), reading-while-listening (Conklin, Alotaibi, Pellicer-Sánchez & Vilkaitė-Lozdienė, 2020), and multimodal reading (Pellicer-Sánchez, Tragant, Conklin, Rodgers, Llanes & Serrano, 2018).

Eye-movement patterns of reading have been used to examine differences between the same individual reading in their L1 vs. L2, as well as differences between individuals varying in L2 proficiency. In a within-subject comparison of bilinguals reading a novel in their L1 (Dutch) vs. L2 (English), Cop, Drieghe, and Duyck (2015) found longer sentence reading times, more fixation counts, less word skipping and shorter saccades during L2 reading (see Cop, Dirix, Drieghe & Duyck, 2017, for similar patterns at the word level in the same dataset). Cop et al. (2015) also found that university-level bilinguals with greater L2 proficiency made fewer fixations per sentence when reading, suggesting that L2 reading efficiency is expected to increase with greater overall L2 proficiency. Similar effects of L2 proficiency, L2 experience or component L2 reading skills on L2 eye-movement behavior during reading have also been reported elsewhere (see

Kuperman, Siegelman, Schroeder, Alexeeva, Acartürk, Amenta, Bertram, Bonandrini, Brysbaert, Chernova, Da Fonseca, Dirix, Duyck, Fella, Frost, Gattei, Kalaitzi, Lõo, Marelli, Nisbet, Papadopoulos, Protopapas, Savo, Shalom, Slioussar, Stein, Sui, Taboh, Tønnesen & Usal, 2022; Nisbet, Bertram, Erlinghagen, Pieczykolan & Kuperman, 2021; Whitford & Titone, 2012, 2017). Together, the results of eye-movement studies indicate that greater proficiency in a language (L1 or L2) is expected to lead to greater reading efficiency when reading texts in the L2.

As indicated above, the eye-tracking literature shows differences between L1 and L2 eye-movement patterns during reading, a finding that aligns with educational research that shows that university-level EALs experience greater reading difficulty compared to native English-speaking students (e.g., Trenkic & Warmington, 2019). However, as far as we are aware, only two eye-tracking studies have been conducted on international EAL populations and none of these has addressed longitudinal (within-participant) change in passage reading. In a study of 38 undergraduate EAL students at a UK university, Bax (2013) found that GLOBAL eye-movements, i.e., eye-movement measures aggregated at the passage level, varied as a function of passage comprehension ability. More recently, Schmidtke and Moro (2021) showed that word-level processing becomes more efficient by the end of the bridging program in a sample of 70 Chinese L1 EALs enrolled in a bridging program. While Bax's (2013) study examined global reading behavior while reading International English Language Testing System (IELTS) reading test texts in EAL students using eye-movements, the study sample was composed of EAL readers ALREADY in undergraduate studies, who had already met the university language proficiency requirements. Moreover, Schmidtke and Moro's (2021) study focused on the contributions of word-level predictors (frequency and length) and component L2 reading skills on the time-course of lexical processing. Therefore, the eye-movement behavior of university-level bridging program students, i.e., students approaching the proficiency level required for university study, during passage reading remains under-researched.

Hypotheses

There are three hypothetical but theoretically plausible profiles for individual differences in reading skill growth among bridging program students. We visualize the three hypotheses in Figure 1 for a generic reading skill measure as an outcome variable (more positive values indicate greater reading skill). The first hypothesis (Panel A.) is that the gap in reading skill between students with high and low initial reading ability diverges over time (e.g., Droop & Verhoeven, 2003). We label this possibility 'divergent change'. The second scenario (Panel B.), which we refer to as 'convergent change', describes the possibility that, despite differences in the hypothetical reading skill measure at the outset of the bridging program, low ability readers gradually close the gap on their more skilled peers (e.g., Raudszus et al., 2021). A third plausible hypothesis (Panel C.) is a scenario in which there is neither divergence nor convergence among readers varying in initial reading ability. This hypothesis would be supported by the observation of parallel developmental paths for readers of all ability levels. We refer to this hypothesis as 'stable change'. Observing stable change could imply that differences in the reading growth trajectories of EALs are set by the beginning of the bridging program, and are maintained across the duration of English language instruction. Such a pattern, if observed, would be predicted by theories of L2 learning that posit an age-related deceleration in L2 learning rate (see Muñoz & Singleton, 2011 for a review). Hence, at a certain point in early adulthood, despite the ability for growth, gross differences in the rate of reading development between EALs are already formed and maintained. We compare the theoretical space of possible growth trajectories (Figure 1) to the patterns present in the experimental data by examining the interactions between testing timepoint and baseline reading ability.

Method

Longitudinal passage comprehension and eye-movement data were collected from three cohorts of students enrolled in an eight-month bridging program for academic English at a Canadian university. Each cohort represents a separate eight-month (28-week) delivery of the same bridging program. For each cohort the same participants were tested at two timepoints: once at the beginning of the program (Time 1: t_1), and once again at the end of the program (Time 2: t_2). Baseline reading profiles and patterns of reading ability growth were evaluated as a function of incoming scores on the READING component of the Academic version of the IELTS test. We examined the possible space of hypotheses in statistical models that examine whether variability in baseline reading skill translates into differences in reading behavior at the start of the bridging program (intercept), and differences in the rate of

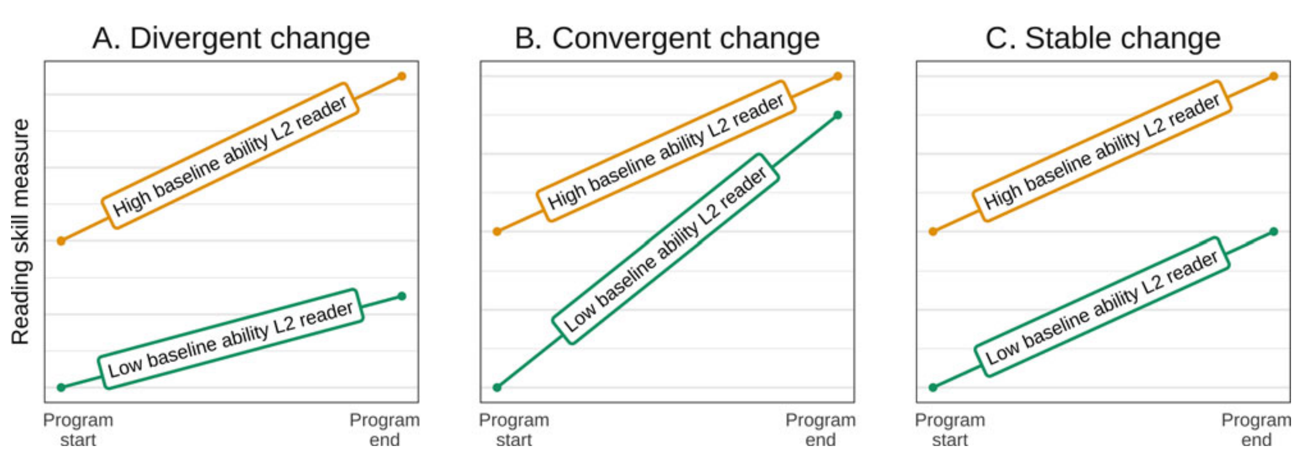

Figure 1. Hypothetical developmental trajectories for bridging program students varying in initial reading skill.

https://doi.org/10.1017/S1366728922000542 Published online by Cambridge University Press

reading development up to the end of the bridging program (growth slope).

Participants

Passage reading data were collected at both timepoints from 430 consenting participants. Data from 25 participants were unusable because of excessive signal loss in the eye-movement recordings. After exclusion of these participants, the data set included passage comprehension scores and associated eye-movement recordings from a combined total of 405 participants (196 female, 205 male, 4 undisclosed) from the 2017–2018 academic year (Cohort 1; data also reported in Schmidtke & Moro, 2021), 2018–2019 (Cohort 2) and 2019–2020 (Cohort 3) sessions. Cohort 1 included 70 participants (34 female), Cohort 2 included 293 participants (149 female and 1 undisclosed) and Cohort 3 included 42 participants (13 female and 3 undisclosed). All participants had normal or corrected-to-normal vision, and none had a diagnosed reading or learning disability. The experiment was administered in English and was part of a larger battery of language tests that were conducted as a quality assurance testing protocol for the bridging program. The eye-tracking and reading comprehension testing component was administered separately from all other tests to minimize fatigue. Participants received course credit for their participation (at both testing timepoints separately). The study was approved by the McMaster University Research Ethics Board (Cohort 1: protocol 2011–165; Cohorts 2 and 3: protocol 2018–239).

English proficiency profile of the participants

While to study at the undergraduate level students must have obtained an overall IELTS score of 6.5 with a minimum of 6.0 in each of the four language components (Reading, Writing, Speaking and Listening), bridging program students must have obtained a minimum overall IELTS score of 5.0. For reference, an overall IELTS 6.5 (corresponding roughly to the uppermost tranche of B2 in the Common European Framework of Reference for Languages, or CEFR) is the language proficiency level typically required for university admission; according to the IELTS band descriptors, a 5.0 is a 'modest' language user, a 6.0 indicates a 'competent' user and a 7.0 indicates a 'good' user. The top score in IELTS is 9.0, i.e., an 'expert' language user.

Participants in our sample entered the program with a median overall IELTS score of 5.5 (range = 5–7), and all were native speakers of Mandarin or Cantonese. A Kruskal-Wallis one-way analysis of variance of ranks confirmed that the differences in median overall IELTS scores between Cohort 1 (median = 5.5), Cohort 2 (median = 5.5) and Cohort 3 (median = 5.5) were not statistically significant; $H(2) = .74$, $p = 0.69$. The average age at the beginning of testing was 18.88 (range = 17.03–22.23). A one-way ANOVA confirmed that the differences in average age across Cohort 1 ($M = 18.88$, $SD = 0.76$), Cohort 2 ($M = 18.89$, $SD = 0.77$) and Cohort 3 ($M = 18.83$, $SD = 0.81$) were not statistically significant; $F(2,400) = 0.13$, $p = 0.87$.

Brief overview of the bridging program

The bridging program provides two terms (28 weeks in total) of English instruction; students are not separated by level; all students complete the same program – same materials, instructional methods, hours, assessments. Modelled on an undergraduate year, the program is a full-time intensive program made up of five courses per term (14 weeks per term). All students complete 10 courses over the eight months, with 21 hours of 'contact' or 'classroom' time per week, most of which (18 of the 21 hours) are done in small classes (15-18 students). Each term, students complete four language development courses and one degree-credit course; one of the four language courses each term is dedicated specifically to reading. The 'reading courses' include readings from a variety of academic disciplines (economics, environmental science, etc.), as well as journalistic pieces; readings increase in difficulty across the program. In these courses, students also learn strategies for dealing with new vocabulary, looking for main ideas, etc., and complete a variety of scaffolded assignments (e.g., paraphrasing, summarizing, critiquing) that culminate in a short literature review. Of course, reading is required across all courses in the program; each course has its own core materials and required readings. For example, the degree-credit course in first term is a linguistics course on the phonetics, phonology and morphology of contemporary English; an undergraduate linguistics textbook is the required reading for this course. All students can avail themselves of additional support through instructional team members.

Procedure

Apparatus

Global eye-movements were recorded using the eye-tracking method and reading comprehension was evaluated with comprehension questions that directly followed the reading of each passage. Eye-movements were recorded using four EyeLink eye-tracking systems (system 1, EyeLink 1000; systems 2, 3, and 4, EyeLink 1000 Plus) manufactured by SR Research Ltd. (Ottawa, Ontario, Canada). The eye-movement camera and a conjoined infrared illuminator were mounted on a desktop beneath the stimulus display. The recording was monocular (right eye). The camera sampled pupil location and pupil size at a rate of 1,000 Hz. A chin support and forehead rest were used to stabilize participants' gross head movements and were placed 60 cm from the screen across all systems. Prior to presentation of the stimuli, in all experiments the eye-tracker was calibrated using a series of nine fixed targets distributed around the display, followed by a nine-point accuracy validation. To ensure that head movements during the oral production of responses did not interfere with the reliability of the eye-movements, experimenters monitored the accuracy of eye-movements between each trial and performed re-calibration and validation steps if necessary. The passages of text were presented on a 17-inch monitor (system 1) and a 24-inch monitor (systems 2, 3 and 4). The NEC MultiSync monitor for system 1 had a resolution of 1,600 x 1,200 pixels, and a refresh rate of 60 Hz. The identical BenQ XL2411Z monitors for systems 2, 3 and 4 had a resolution of 1,920 x 1,080 pixels, and a refresh rate of 144 Hz. Across all systems the passages were presented in 15-point Courier New fixed-width font: one degree of visual angle was subtended by about 3.3 characters. Texts were also presented across the same number of lines, with the same number of line breaks, across all systems.

The eye-tracking system was included as a co-variate in all analyses; it was not found to be a statistically significant predictor of any of the analyzed outcome measures. In addition, to ensure that results were not impacted by the hardware specifications of system 1 (a smaller screen, lower resolution and longer refresh rate), all analyses (see *Statistical method*) were conducted with data obtained from system 1 removed (3% of total number of trials for passage comprehension data and 2% of the total number of trials for the eye-movement data). The model outputs indicated that the critical patterns in the reduced data do not differ from those obtained from the complete data source.

Materials

Participants silently read six stories (stories 3-8) of increasing complexity taken from the Gray Oral Reading Tests - Fifth Edition (GORT-5, Wiederholt & Bryant, 2012). Cohort 1 read passages from level 9 of GORT-5, but these data were removed since the median comprehension accuracy for this passage was 0%. Each passage of text was preceded by a drift correction, which used a fixation point positioned 20 pixels to the left of the beginning of the first line of the story. Stories were presented 100 pixels away from the left edge of the screen, and in the middle of the vertical dimension of the screen. Participants were instructed to press a button when they had finished reading each passage of text, and the text remained on the screen until the button was pressed. After reading each passage of text, participants answered five comprehension questions each presented on a separate screen. The format of the GORT-5 comprehension questions is open-ended, allowing participants to freely produce a response. Responses were recorded by the experimenter. Once participants answered a question, they were instructed to press a button for the next question to appear until all five questions had been answered. Answers were spoken aloud by each participant and were recorded by the experimenter. After answers to all five questions were recorded, the next trial began.

Each story and each question occupied exactly one screen; the longest text occupied 11 lines. Stories were presented in a fixed order of complexity (from least difficult to most difficult). The GORT-5 testing kit contains two versions of the test: Form A and Form B. At t_1, half of the participants read passages from Form A, and the remaining half of participants saw passages of equivalent complexity from Form B. Forms were randomly assigned. Story assignment was counterbalanced across timepoints: participants who read Form A at t_1 were presented with passages from Form B at t_2, and vice-versa. The testing session lasted no longer than 60 minutes with most participants completing the experiment within 30 minutes (including breaks). Table 2 provides descriptive statistics for the text characteristics of the GORT-5 passages. Absolute Spearman rank correlation coefficients (ρ) between complexity and characteristics were high for each form (min = 0.77; max = 0.94), confirming that text complexity measures are strongly related to the GORT-5 passage complexity number.

Test reliability

Parallel forms reliability was examined by comparing the differences in means and variances of total correct responses between GORT-5 Form A and Form B at each timepoint (Gulliksen, 1950). An indepennt two-sample t-test confirmed that the difference in average scores between Form A ($M = 14.6$, $SD = 3.94$) and Form B ($M = 14.13$, $SD = 4.19$) at t_1 was not statistically reliable [$t(392.5) = 1.15$; $p = 0.25$]. The difference in mean accuracy between Form A ($M = 15.77$, $SD = 4.1$) and Form B ($M = 15.1$, $SD = 4.13$) at t_2 was also not statistically reliable [$t(399.36) = 1.54$; $p = 0.12$]. Differences in variances across lists were examined using F-tests, which confirmed that the differences in variances in the total number of correct responses between Form A and Form B were not statistically significant at both t_1[$F = 0.89$, $p = 0.4$] and t_2[$F = 0.98$, $p = 0.91$]. These results indicate that Form A and Form B elicited consistent results at each timepoint. Cronbach's alpha for the GORT-5 comprehension data was computed using the *ltm* package in R (Rizopoulos, 2006), with the number of bootstrap samples set to 1,000. Cronbach's alpha was .65, 95% CI [.60, .70] for Form A and .68, 95% CI [.63, .72] for Form B. The participant-level reliability for silent reading rate was obtained for each form using the Intra-class Correlation Coefficient (ICC), which estimated the degree of participant agreement in reading rate across the 6 texts. The ICC was 0.77 for Form A and 0.86 for Form B, demonstrating high agreement among participants across texts for each form.

Dependent variables

Scores of the GORT-5 passage comprehension test were used as a measure of reading comprehension. These scores ranged between 0-5 and were entered into analyses as an accuracy measure, i.e.,

Table 2. Descriptive statistics for the passages. Standard deviations provided in parentheses. Spearman's rho reported for the relationship between passage complexity and each measure.

		Passage						
	Measure	3	4	5	6	7	8	ρ
Form A	Number of words	52	85	106	100	107	130	0.94
	Mean sentence length (words)	7.43 (1.9)	10.62 (2.3)	15.14 (5.21)	11.11 (3.79)	17.83 (4.54)	18.57 (2.88)	0.94
	Mean word length (characters)	3.48 (1.77)	3.81 (1.55)	3.79 (1.68)	4.21 (2.02)	4.46 (2.17)	4.89 (2.51)	0.94
	Mean log word frequency (content words)	4.5 (0.73)	3.68 (0.81)	3.67 (1.03)	3.81 (0.79)	3.27 (1.08)	3.31 (1.19)	−0.77
Form B	Number of words	52	82	102	103	98	125	0.83
	Mean sentence length (words)	8.67 (2.25)	9.11 (1.69)	12.75 (5.14)	12.88 (3.55)	14 (5.16)	13.89 (4.92)	0.94
	Mean word length (characters)	3.77 (1.64)	3.79 (1.37)	3.97 (2.12)	4.06 (1.81)	4.84 (2.18)	4.59 (2.31)	0.94
	Mean log word frequency (content words)	4 (0.96)	4.08 (0.62)	3.68 (0.83)	3.66 (0.85)	2.95 (1)	3.25 (0.95)	−0.89

Note. Word frequencies obtained from the SUBTLEX-US corpus (Brysbaert & New, 2009).

the proportion of correct responses per passage. Five global eye-movement measures were analyzed, where each measure quantifies an aspect of text processing fluidity for an individual reading a passage of text: total number of fixations, total number of regressions (defined as the number of regressions from a word to earlier parts of the text), total number of skips, total fixation duration (ms), and silent reading rate (words per minute). Words per minute was computed as the total number of words per passage divided by the total reading duration per passage (in minutes). Total fixation duration, total number of fixations, total number of regressions, and total number of skips were normalized by the number of words per passage. The eye-movement measures, in addition to silent reading rate, are described below.

Skipping

Word skipping is an early measure that captures automatic processes associated with text reading (Conklin & Pellicer-Sánchez, 2016). Skipped words are processed prior to or after the skip, and the likelihood of word skipping is increased for shorter, more frequent words and words that are more predictable in context (see Rayner, 2009, for a comprehensive review). These findings are reflected at the global level: skipping counts are lower for more complex passages of text (Rayner, Pollatsek, Ashby & Clifton, 2012). Moreover, more skilled readers, and readers with greater reading experience and vocabulary knowledge, tend to skip words more often (Ashby, Rayner & Clifton, 2005; Kuperman, Matsuki & Van Dyke, 2018).

Regressions

Regressions provide a measure of how many times a reader returns to an earlier part of a text. Eye-movement research has established that regressions are linked to disruptions in higher order text processing (Frazier & Rayner, 1982; Inhoff, Weger & Radach, 2005; Joseph & Liversedge, 2013; Meseguer, Carreiras & Clifton, 2002; Rayner, Chace, Slattery & Ashby, 2006; Vorstius, Radach, Mayer & Lonigan, 2013). Readers tend to revisit earlier parts of a text when experiencing difficulty integrating the meaning of individual words with the overall meaning of the passage, or when experiencing difficulty in syntactic processing. In short, readers who make more regressions per text experience greater comprehension difficulties at the sentence or passage level.

Fixation-based measures

Fixation counts generally increase when readers engage with more challenging texts (e.g., Rayner et al., 2006). It has also been shown that the number of fixations increases when a reader encounters an unfamiliar word, or a word that has lexical properties that make it difficult to process, such as longer or infrequent words (e.g., Schmidtke & Moro, 2021). Less skilled readers also make more fixations per passage (Spichtig, Pascoe, Ferrara & Vorstius, 2017). Fixation counts correlate with total fixation duration (Godfroid, 2020): more fixations are associated with longer overall fixation durations.

Both words per minute and total fixation duration (ms) were log-transformed for analysis (see *Statistical method*). Table 3 displays the means, standard deviations, and ranges of all dependent measures, and their transformed values (where relevant), broken down by Timepoint.

Independent variables

Baseline reading ability

Scores on the Reading component of the Academic version of the IELTS exam were used as the measure of initial reading ability in the bridging program (henceforth referred to as Baseline reading ability). This approach avoids the artifactual statistical interpretation of patterns in the data resulting from regression to the mean, an issue that would be imposed by grouping students based on their performance on the dependent measure itself (Barnett, van der Pols & Dobson, 2005; Phillips, Norris, Osmond & Maynard, 2002; Protopapas, Sideridis, Mouzaki & Simos, 2011). A histogram of Baseline reading scores is provided in Figure 2. Distributional characteristics of Baseline reading ability for the entire sample and for individual cohorts are provided in Table 4. A Kruskal-Wallis one-way analysis of variance of ranks confirmed that the differences between median IELTS Reading subtest scores across cohorts were not statistically significant; $H(2) = 0.038$, $p = 0.98$. An independent two-sample Mann-Whitney U-test confirmed that the association between the assignment of GORT-5 forms at t_1 (either Form A or Form B) and Baseline reading ability was not statistically significant; $U = 21692$; $p = 0.28$. IELTS scores are considered valid by university admissions if the test was taken within 2 years of application to the bridging program. The length of time between IELTS test date and the date of participation in the experiment was included in our analyses and it did not predict results; this variable is therefore not discussed further.

In addition to Baseline reading ability, Timepoint was used as an index of longitudinal change (two levels: t_1 and t_2). Passage complexity was included as a control variable and was defined as the ordinal number of the text in the GORT-5 test kit (3–8): the higher the number, the more complex the passage (see *Materials*).

Statistical method

A series of (generalized) linear mixed-effects multiple regression models (Baayen, Davidson & Bates, 2008; Jaeger, 2008; Pinheiro & Bates, 2000) was fitted to the reading data. Passage comprehension data were fitted using a binomial underlying distribution for the proportion of correct responses per passage. Eye-movement measures were fitted using a Gaussian underlying distribution. Data analysis was performed using the *lme4* package (using the BOBYQA algorithm for optimization, Bates, Mächler, Bolker & Walker, 2015) in the R statistical environment (Version 4.1.2; R Core Team, 2021). The *p*-values for model fits were obtained using the *lmerTest* package (Kuznetsova, Brockhoff & Christensen, 2017), which uses Satterthwaite's degrees of freedom method.

Regression models were fitted to the log-transformed values of total fixation duration (in ms) and silent reading rate (in wpm) to account for the skewness of their distributions. We assessed differences between Timepoints using successive repeated difference contrast coding, implemented using the *contr.sdif()* function in the *MASS* package in R (Venables & Ripley, 2002). Baseline reading ability and Passage complexity were included in models as continuous variables. Baseline reading ability and Passage complexity were scaled to ensure comparability in regression analyses. Models included the main effects of Timepoint and Baseline reading ability, and the interactions between these variables. The possible nonlinear nature of the effect of Baseline reading ability was explored by fitting quadratic and cubic polynomial curves to the data using the *poly* function in R. Step-wise model comparisons

Table 3. Descriptive statistics of passage comprehension scores and eye-movement measures broken-down by Timepoint.

	Original			Transformed		
Measure	*M*	*SD*	*range*	*M*	*SD*	*range*
Passage comprehension (%)						
t_1	50.25	27.24	0–100			
t_2	53.84	27.43	0–100			
Silent reading rate (words per minute)						
t_1	120	47	36–272	4.79	3.85	4–6
t_2	126	47	40–277	4.84	3.85	4–6
Total number of fixations (per word)						
t_1	3.059	1.082	1.278–8.711			
t_2	2.853	0.929	1.279–7.583			
Total number of regressions (per word)						
t_1	0.187	0.062	0.023–0.4			
t_2	0.185	0.064	0.014–0.444			
Total number of skips (per word)						
t_1	0.159	0.079	0–03			
t_2	0.167	0.081	0–13			
Total fixation duration (per word, ms)						
t_1	587	254	220–1650	6.38	5.54	5–7
t_2	554	229	217–1500	6.32	5.43	5–7

showed that neither non-linear function forms improved model fit. Interactions between Passage complexity and Baseline reading ability were also included to examine whether performance in passages of varying complexity was contingent on reading ability. All models included random intercepts for trial (the passage ID) and by-participant random intercepts with random slopes for Timepoint. The by-participant random effects term takes into account the correlated observations that characterize longitudinal data.

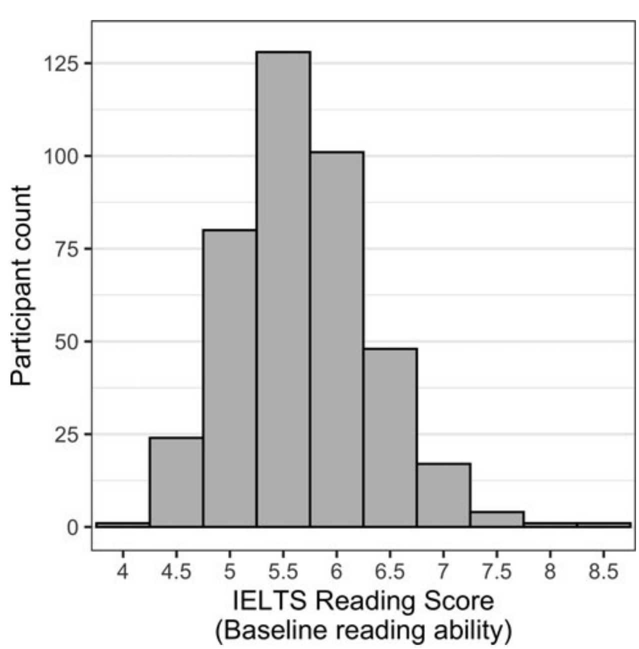

Figure 2. Histogram of the distribution of IELTS Reading scores as measure of Baseline reading ability.

The following modelling procedure was applied to each outcome measure. First, a full model was fitted that included the Timepoint-by-Baseline reading ability interaction. This interaction assessed the critical hypothesis of interest, i.e., longitudinal (slope) differences as a function of baseline reading ability. The generic formula for this regression model was:

DV ~ Timepoint * Baseline reading ability + Passage complexity * Baseline reading ability + (1 | Passage ID) + (1 + Timepoint | Participant ID),

where DV is the dependent variable. Second, for each measure, an unconditional model was fitted that included Timepoint and Baseline reading ability as main effects:

DV ~ Timepoint + Baseline reading ability + Passage complexity * Baseline reading ability + (1 | Passage ID) + (1 + Timepoint | Participant ID).

Chi-square goodness-of-fit tests were then applied to compare AIC values of the full model with the unconditional model. The results of the best fitting models are presented as the final results. Further model criticism involved refitting models after removing outliers from all data sets. This was conducted by excluding absolute standardized residuals exceeding 2.5 standard deviations (Baayen & Milin, 2010). The false discovery rate procedure (Benjamini & Hochberg, 1995) was applied in addition to the rule-of-thumb criterion to control for the inflated Type I error rate resulting from the evaluation of the same effects of interest

https://doi.org/10.1017/S1366728922000542 Published online by Cambridge University Press

Table 4. Descriptive statistics of Baseline reading ability as measure by IELTS Reading subtest scores.

	n	Mean	*SD*	Minimum	25%	Median	75%	Maximum
Entire sample	405	5.68	0.66	4	5	5.5	6	8.5
Cohort 1	70	5.69	0.6	4.5	5.5	5.5	6	7.5
Cohort 2	293	5.67	0.69	4	5	5.5	6	8.5
Cohort 3	42	5.69	0.57	5	5.5	5.5	6	7

across six separate models (Von der Malsburg & Angele, 2017). Thus, any effect was deemed reliable if it was significant after the application of the false discovery rate correction or if a significant effect was observed in at least two out of six of the analyzed measures (probability of observing 2 out of 6 effects with alpha set at .05 is = 0.033). Effect estimates were extracted from models in R using the *effects* package (Fox & Weisberg, 2019). All plots were generated using the *ggplot2* package (Wickham, 2016).

Power analysis

A power analysis was conducted to ensure that the results could be expected to replicate (Brysbaert & Stevens, 2018; Brysbaert, 2021). Simulations were used to estimate power with the *simr* package for linear-mixed models (Green & MacLeod, 2016). The simulations were based on data collected from Cohort 1 (2017–18; $n = 70$), using the *powerCurve()* function to estimate the number of subjects required to reach 80% power for each outcome variable. For each measure, a model was fitted as described in *Statistical method*, where 100 random samples were selected for a range of simulated sample sizes from 100 to 400 participants with step-wise increments of 50 participants. For all outcome measures and for each effect of interest (main effects of Timepoint and Baseline reading ability, and their interaction) the effect size was set at $d = .4$. This effect size matches the smallest effect size for change in words per minute from prior studies listed in Table 1, and is also the effect size that is recommended for power analyses in novel psychological research (Brysbaert, 2019a). The results indicated a sample size of at least 100 would be needed to achieve > 80% power for a main effect of Time, a main effect of Baseline reading ability, or an interaction between both variables for all of the outcome variables with a effect size of $d = .4$. Data were collected from over 400 participants to account for potential participant attrition, data loss due to poor comprehension, experimenter error or equipment failure.

Results and discussion

The initial data set contained 4,860 passage readings (405 participants x 6 passages x 2 timepoints). We removed 124 (2.6%) passage readings due to signal loss, skipping and excessive blinking. We further removed passage readings for which reading speed was in the bottom 1% of words per minute (below 36 wpm at t_1 and below 40 wpm at t_2) and in the top 1% of words per minute (above 272 wpm at t_1 and above 277 wpm at t_2). The resulting data set for the analysis of passage comprehension contained 4,640 data points ($t_1 = 2{,}317$; $t_2 = 2{,}323$). For the eye-movement data, we analyzed readings for which participants exhibited adequate comprehension (see also Spichtig et al., 2017). The analysis of the eye-movement record is therefore based on the subset of data for which passage comprehension reached a 60% threshold of accuracy (the median comprehension score). This led to the removal of 2100 (45.3%) passage readings (including the removal of all data from 2 participants). The final eye-movement data set contained 2,540 ($t_1 = 1{,}182$; $t_2 = 1{,}358$) valid passage readings from 403 participants.

In the following sections we provide the results of the regression analyses of passage comprehension and eye-movements during reading. We did not observe any statistically significant interaction effects between Timepoint and Baseline reading ability. Model comparisons showed the interaction between Timepoint and Baseline reading ability did not significantly improve model fit for all outcome measures (all $\chi^2 < 2.5$, $p > 0.05$). Tables S1.1.–S1.6 in the supplementary materials (Supplementary Materials) provide full specifications of each model. The effect sizes from regression models (plotted in Figures 3 and 4) show the fitted partial effects of words per minute and fixation durations (ms) after the back-transformation of log values to their original scales. Plots of raw trends are found in Figure S2.1 in the supplementary materials (Supplementary Materials).

Main effects of Timepoint

Focusing on the main effect of Timepoint, at t_2 compared to t_1, students made significantly more skips per word [$\hat{\beta} = 0.006$; $SE = 0.003$; $t = 2.271$; $p = 0.02$], fewer fixations per word [$\hat{\beta} = -0.159$; $SE = 0.032$; $t = -4.92$; $p < 0.001$], and had significantly faster reading rates [$\hat{\beta} = 0.049$; $SE = 0.013$; $t = 3.666$; $p < 0.001$] and shorter total fixation durations per word [$\hat{\beta} = -0.049$; $SE = 0.013$; $t = -3.666$; $p < 0.001$]. The main effect of Timepoint on regressions per word was not statistically significant [$\hat{\beta} = -0.001$; $SE = 0.002$; $t = -0.379$; $p = 0.7$]. There was also a main effect of Timepoint on passage comprehension [$\hat{\beta} = 0.177$; $SE = 0.032$; $z = 5.469$; $p < 0.001$], indicating that, on average, students became better at comprehending passages over the duration of the bridging program. Overall, these results indicate that there were significant within-participant gains in passage comprehension and in all but one measure of passage reading efficiency. Model-based estimates of average change over time are presented in Table 5.

Main effects of Baseline reading ability

We observed statistically significant main effects of Baseline reading ability on all outcome measures: passage comprehension [$\hat{\beta} = 0.147$; $SE = 0.024$; $z = 6.112$; $p < 0.001$], silent reading rate [$\hat{\beta} = 0.065$; $SE = 0.013$; $t = 5.115$; $p < 0.001$], number of fixations per word [$\hat{\beta} = -0.122$; $SE = 0.03$; $t = -4.109$; $p < 0.001$], number of regressions per word [$\hat{\beta} = -0.005$; $SE = 0.002$; $t = -2.698$; $p = 0.008$], number of skips per word [$\hat{\beta} = 0.011$; $SE = 0.003$; $t = 4.285$; $p < 0.001$], and total fixation duration per word [$\hat{\beta} = -0.065$; $SE = 0.013$; $t = -5.115$; $p < 0.001$]. The effect sizes

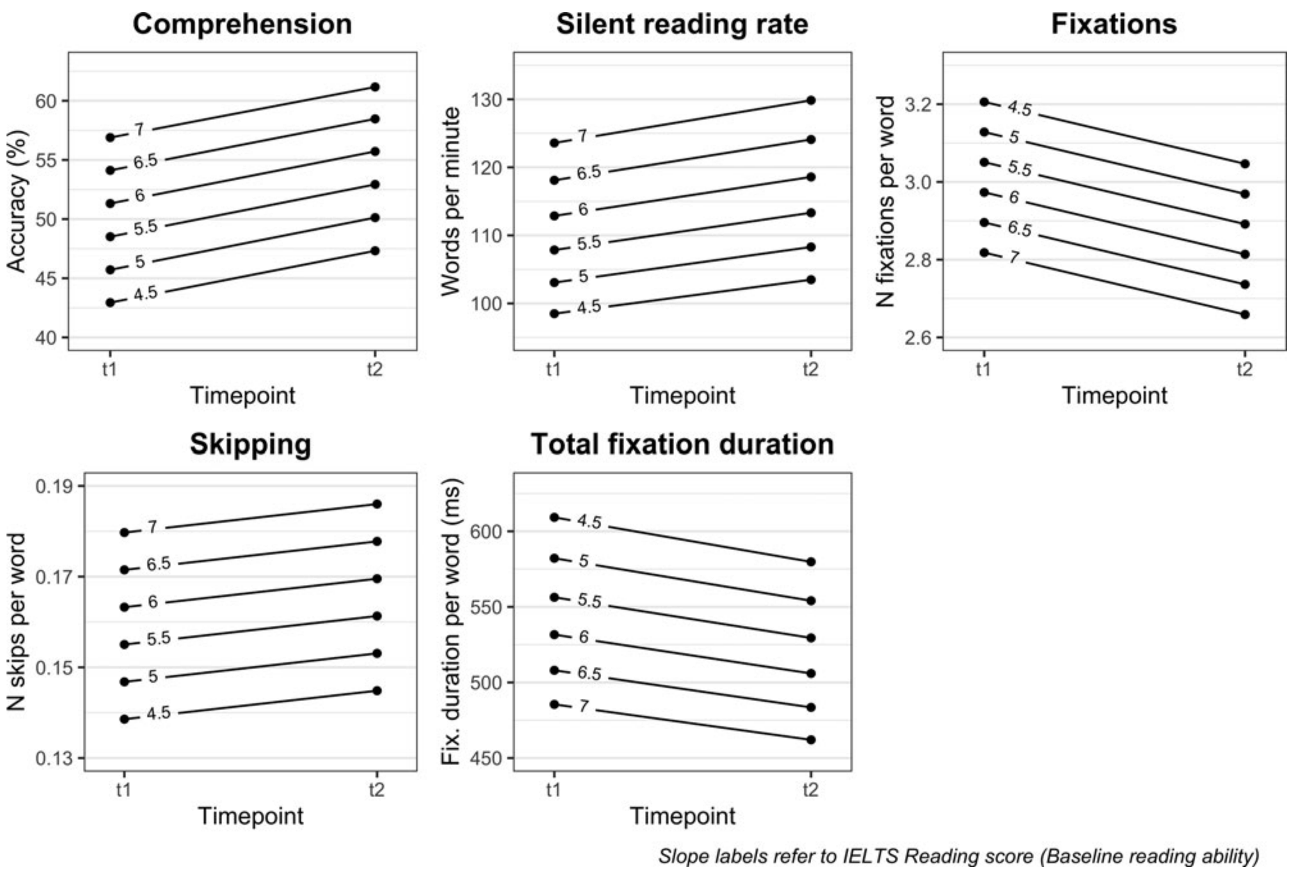

Figure 3. Partial main effects of Timepoint and Baseline reading ability on reading measures.

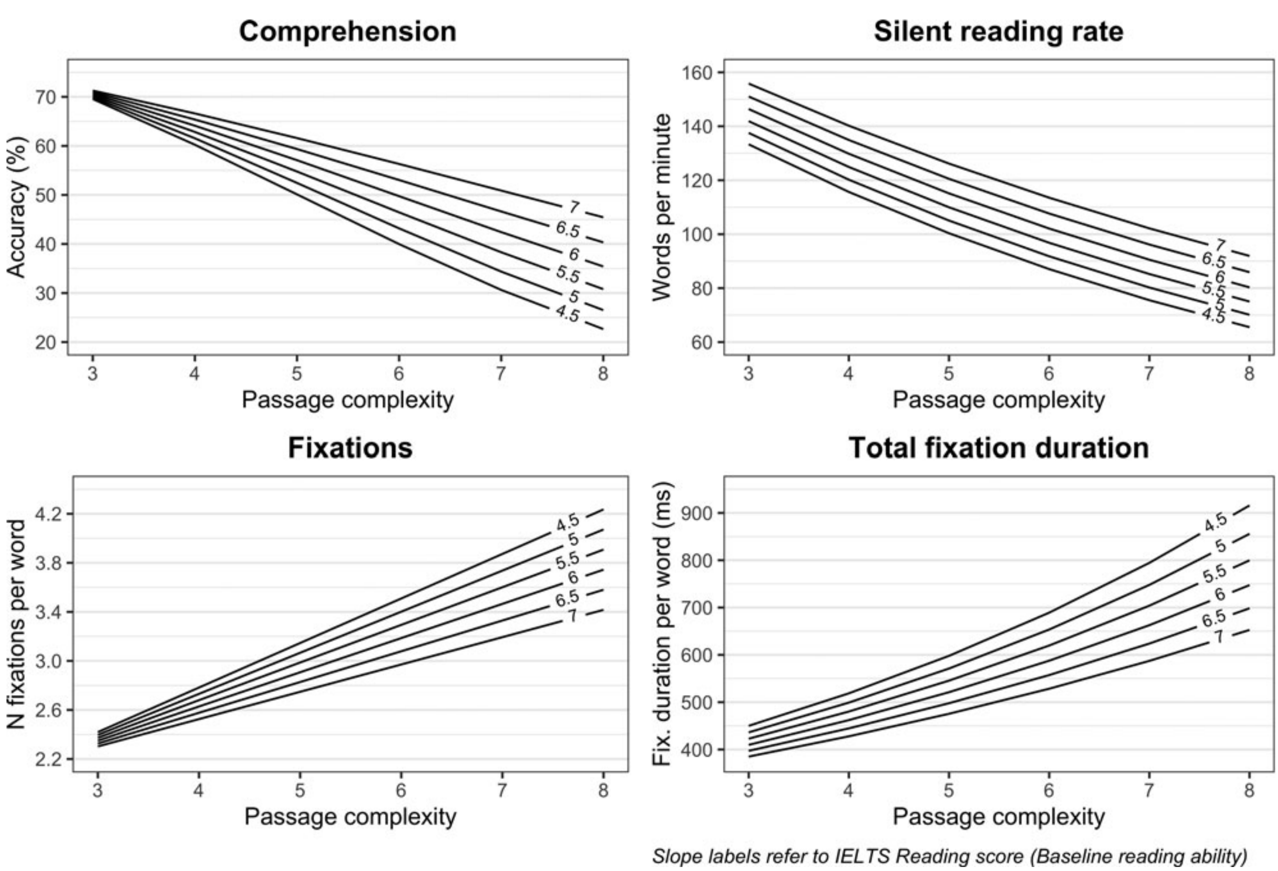

Figure 4. Partial effects of Passage complexity modulated by Baseline reading ability on reading measures.

https://doi.org/10.1017/S1366728922000542 Published online by Cambridge University Press

Table 5. Model-based estimates of timepoint averages and predicted growth over time. Growth expressed as absolute change (Δ) and relative change.

Measure	t_1	t_2	Δ	Relative change
Passage comprehension (%)	50	54	4	8%
Silent reading rate (wpm)	111	116	6	5.07%
Total number of fixations (per word)	3.012	2.853	−0.16	−5.29%
Total number of regressions (per word)	0.186	0.185	-0.001^{ns}	−0.47%
Total number of skips (per word)	0.159	0.165	0.006	3.95%
Total fixation duration (per word, ms)	545	519	−26	−4.83%

^{ns}indicates non-significant effect.

indicate that a .5 increase in the IELTS Reading subtest is predicted to lead to an increase in 2.78% in passage comprehension accuracy, an increase of 5 words per minute, a decrease in 0.08 fixations per word, a decrease in −0.004 regressions per word, a decrease in 0.01 skips per word, and a reduction of 24 ms in fixation durations per word.

Figure 3 shows the partial main effects of Timepoint and Baseline reading ability. The figure displays the predicted slopes for values of the IELTS reading subtest ranging from 4.5 to 7 (representing the range of IELTS scores for 98.3% of the total number of observations for the comprehension data and 98.2% of the eye-movement data). In sum, the results confirm that incoming reading ability is a predictor of eye-movements and reading comprehension. The results indicate differences in intercepts between Baseline reading ability at t_1, and that their slopes over time follow a parallel trajectory, resulting in near-identical differences at t_2. The observation of parallel developmental paths for all baseline reading levels therefore lends support to the stability pattern of growth in reading skill.

Interaction effects between Baseline reading ability and Text complexity

Across all reading measures there were statistically significant main effects of passage complexity: as expected, more difficult passages gave rise to poorer comprehension scores and elicited slower reading rates (wpm), fewer skips, more fixations and regressions, and slower total reading times (see Tables S1.1.–S1.6. for model estimates for these main effects). There were also significant interactions between Passage complexity and Baseline reading ability in passage comprehension [$\hat{\beta} = 0.086$; $SE = 0.015$; $t = 5.913$; $p < 0.001$], words per minute [$\hat{\beta} = 0.016$; $SE = 0.004$; $t = 3.955$; $p < 0.001$], number of fixations per word [$\hat{\beta} = -0.063$; $SE = 0.01$; $t = -6.01$; $p < 0.001$], and total fixation duration per word [$\hat{\beta} = -0.016$; $SE = 0.004$; $t = -3.955$; $p < 0.001$]. These interactions had a common pattern. Namely, an increase in text complexity affected the reading behavior of a weaker EAL reader more than a stronger EAL reader. As shown in Figure 4, when reading complex passages, participants with lower reading ability comprehended less, read fewer words per minute, made longer fixations, and executed more fixations compared to participants with a relatively higher baseline reading ability. Therefore, readers entering the bridging program with lower IELTS Reading subtest scores are expected to find more complex texts particularly challenging to read. Importantly, there were no statistically significant three-way Timepoint x Baseline reading ability x Passage complexity interactions, indicating that these baseline reading differences across passages of varying complexity did not change over time.

The relationship between reading rate and passage comprehension

We examined the relationship between comprehension and reading rate in the entire data set. The correlation between silent reading rate (wpm) and comprehension score (%) was weak but positive at both $t_1[r = 0.17]$ and $t_2[r = 0.19]$, both $ps < 0.001$. In sum, reading speed was associated with reading comprehension in such a way that passages that were read faster were also comprehended more accurately, and this relationship was slightly stronger at t_2. We discuss this result in the General Discussion.

General discussion

The aim of this study was to gain a basic understanding of passage reading ability of EAL students during the course of a pre-university English bridging program. We investigated two interrelated issues regarding the development of L2 reading ability in this population of EALs. Namely, we examined whether individual differences in reading ability measured at the outset of the bridging program was linked to (i) early differences in global reading skills, and (ii) different trajectories of change in global reading ability.

Longitudinal changes in comprehension and eye-movements during passage reading

Changes in eye-movements and reading comprehension signaled a shift toward greater reading proficiency among this sample of EAL bridging program students. Within-participant growth over the 28-week period was statistically significant for silent reading rate, skipping rate, number of fixations per word, total fixation duration per word and reading comprehension. Average growth was fairly minimal for silent reading rate (Table 3) and is in accordance with the smaller effect sizes reported by Beglar et al. (2012) and McLean and Rouault (2017) (Table 1) for similar 28- to 30-week study duration. Despite the significant change, by the end of the bridging program, students reached an estimated speed of 103–130 words per minute. Based on a comparison with estimates obtained from Brysbaert's (2019b) meta-analysis of reading rates, we can approximate that by the end of the bridging program, the average EAL student is able to cover roughly half as many words per minute compared to the average L1 English reader (116 wpm versus 238 wpm).

Notably, we did not observe a statistically significant reduction in regressive saccades (see also Schmidtke & Moro, 2021). Regressions to earlier parts of text are linked with disruptions to higher level processing of passages (Cook & Wei, 2017, 2019), reflecting processes that occur downstream in global discourse processing, such as the re-evaluation and integration of information across multiple sentences and paragraphs (e.g., Frazier & Rayner, 1982; Rayner et al., 2006). However, as would

https://doi.org/10.1017/S1366728922000542 Published online by Cambridge University Press

be expected, more complex passages elicited a greater number of regressive saccades. Together, these findings might suggest that EALs experienced more difficulties integrating and validating information across multiple sentences, and that the development of reading strategies that contribute to higher level text comprehension remained static over time.

The presence of growth in fixation counts, fixation durations, and skipping, in the absence of change in regression rates, suggests that shifts in reading efficiency are not driven by changes to oculomotor behavior reflecting high level semantic processes (see Schmidtke & Moro, 2021 for a similar finding in an examination of the lexical processing timecourse). Therefore, it may be that gains in reading speed are linked more closely to increases in word skipping (see below) or shorter word fixation times within the first pass of reading. This finding is supported by the correlations between the individual growth curves across the linear-mixed effects models. The strength of the correlation between the by-participant random slopes for timepoint for the number of regressions per word and silent reading rate was estimated at $r = 0.06$. On the other hand, the same correlation between silent reading rate and number of fixations per word was $r = -0.82$, and between silent reading rate and skipping rate per word was $r = 0.48$. In other words, longitudinal growth in silent reading rate is more strongly associated with reductions in fixations and, to a lesser extent, with increases in skipping rates, than to reductions in regressive saccades (refer to Figure S3.1 in the Supplementary Materials for a correlation matrix showing the associations between the individual growth curves of all outcome measures). Overall, our findings suggest that increased experience with language is associated with faster text processing speed (e.g., Cop et al., 2015; Gordon, Moore, Choi, Hoedemaker & Lowder, 2020; Spichtig et al., 2017; Whitford & Titone, 2012), particularly in eye-movement fixations associated with the earlier stages of word identification, rather than with the integration and evaluation of information within the wider discourse context.

A novel aspect of our study is the examination of skipping rates in EAL reading behavior. Overall, EAL students tended to skip less when reading more difficult passages of text. Skipping occurs when the reader acquires information about upcoming words from parafoveal vision, which causes the eyes to skip over the target word and land on the following word (Rayner, 1998, 2009). One textual source of information known to impact skipping (see Rayner, 2009 and references therein) is word length. It is therefore likely that EAL students tend to skip fewer words in more complex passages because they contain longer words (Table 2). In addition, we found that students tended to skip more words per passage by the end of the program, and that skipping growth occurred at the same rate for all passages' complexity levels. This finding suggests that the boost in silent reading rate stems not only from shorter and fewer fixations on words, but also from the ability to identify words (both long and short) in the parafovea more efficiently.

Stable change

Turning to the question of skill-based differences in growth trajectories, comparisons of slopes indicated that there were no significant differences in growth trajectories between readers of varying baseline ability across all reading measures (Figure 3). Instead, estimated slopes indicated parallel growth trajectories for readers of all initial ability. This result is consistent with the stable change hypothesis (Panel C. Figure 1): readers who are weaker at the outset neither close the gap (convergent change) nor fall behind (divergent change) stronger readers in eye-movement measures of reading speed or passage comprehension. Drilling down to the participant level, the stability pattern is also supported by the correlations between the random intercept (individual initial starting point at t_1) and the random slope for Timepoint (individual rate of change) obtained for each linear mixed-effect model (see Supplementary Materials S1). Pearson correlation coefficients ranged from -0.22 to 0.07 across all models, indicating a weak relationship between the starting point of each eye-movement measure and subsequent growth, i.e., stability.

We can offer at least three possible reasons for the stable change result. The first is that the gaps in English reading ability that were formed earlier in participants' developmental trajectory had become set by the time the students entered the bridging program. Second, it is possible that divergent change was not observed because the bridging program somehow served the weakest readers effectively enough (even though all students completed the same program), counteracting a potentially widening gap with respect to the strongest readers, but not as much as would be needed to enable them to catch up to the stronger readers (convergent change). Third, it is also possible that the length of the program is simply insufficient to capture either a pattern of divergence or convergence. We call for future research to address these important issues (see *Implications, limitations and future directions*).

The influence of baseline reading ability

The main effects of Baseline reading ability indicate that differences between students varying in initial ability were present at t_1 and, as discussed above, were sustained at t_2. Students with relatively higher IELTS reading scores began the bridging program with a clear 'head-start' in terms of comprehension and reading fluency. To give a sense of scale, we offer a comparison between a student with an incoming IELTS reading subtest score of 5.0 and one with a score of 7.0, although it's worth noting that a reading subtest score of 7.0 is not even required for direct undergraduate entry. Compared to a student who starts the program with an IELTS Reading score of 5.0, a student with an IELTS Reading score of 7.0 is expected to be 22% stronger at reading comprehension, cover 20% more words per minute, spend 17% less time fixating on words, make 10% fewer fixations, execute 8% fewer regressions to earlier parts of texts and skip 22% more words. These differences reflect the predictive power of IELTS Reading scores for eye-movement measures of reading (cf. Bax, 2013), and for passage comprehension, among EAL bridging program students.

Reading speed and reading comprehension

We found that faster silent reading rates were associated with higher passage comprehension of the same texts, which is consistent with a number of L1 and L2 reading studies (Fuchs, Fuchs, Hosp & Jenkins, 2001; Jenkins, Fuchs, van den Broek, Espin & Deno, 2003; Hebert, 2017; Priya & Wagner, 2009; Roehrig, Petscher, Nettles, Hudson & Torgesen, 2008; McLean & Rouault, 2017). An explanation for the link between successful construction of meaning from texts and faster reading speed is that efficient word identification abilities free up cognitive resources that are responsible for higher-level processing involved in passage reading, such as the mental organization of the semantic structuring of texts (LaBerge & Samuels, 1974; Stanovich, 2000). This also supports Carver's (1990) model of OPTIMAL READING EFFICIENCY: students are able to maintain reading speed without loss in comprehension. Finally, the positive relationship between reading rate and reading comprehension (see also Hirai, 1999) was observed at both timepoints, and was slightly stronger at t_2. This may suggest that EALs in the study achieved greater optimal reading efficiency over time: the factors underlying reading speed and comprehension become more tightly linked.

Implications, limitations and future directions

We found evidence of a gap between readers of varying incoming reading ability at the beginning of the bridging program and that this gap remained stable for the remainder of the program. These results imply that students entering the bridging program with relatively lower IELTS Reading scores are at a disadvantage over other students from the outset, and that this gap is likely to persist at least until the end of the program. An obvious implication for bridging program educators is to use the incoming reading profiles afforded by the IELTS Reading scores as an early warning system. The IELTS Reading scores could be used as a metric for allocating additional reading support.

It is worth recalling that all the EALs in this study are L1 speakers of Mandarin and Cantonese who were proficient in a logographic (morphosyllabic) writing system before learning to read English orthography. As such, we cannot determine the role played by knowledge of the L1 writing system on reading measures. Given the relative linguistic homogeneity of the bilingual cohort in this study, and the specific nature of any positive transfer between L1 and L2 orthographic learning (Geva, Xi, Massey-Garrison & Mak, 2019; Keung & Ho, 2009; Koda, 2008; Wang, Perfetti & Liu, 2005; Verhoeven, Perfetti & Pugh, 2019), it is prudent to be cautious about generalizing the developmental pattern observed here to EALs at a similar stage of English proficiency with different L1 profiles.

Future research might address additional limitations in our study. At present we also do not know whether the reading gap registered at the beginning of the bridging program translates into persistent long-term differences in academic attainment or reading skill post-bridging program completion. It would be valuable to pair data obtained from students in this sample with grade point average data and reading data during undergraduate study.

Conclusion

Our results shed new light on the challenges facing bridging program students when reading passages of text in English. Our findings show that overall improvements in passage reading fluency and passage comprehension are to be expected during a bridging program, and that these improvements appear to be driven by reductions in fixations and increased skipping rates, but not regressions. Students with relatively lower incoming IELTS Reading scores tend to exhibit weaker passage comprehension, read fewer words per minute, make more fixations, make more regressions, tend to skip fewer words, and spend more time processing words. These differences persist over the duration of the program. It is therefore important to direct targeted reading support to students within the lower range of IELTS Reading scores to help them develop reading strategies that can support learning during undergraduate studies.

Data availability statement

The data that support the findings of this study are not available due to restrictions in the ethics protocol. R code used for analysis is available upon request.

Acknowledgments. Daniel Schmidtke and Anna L. Moro are affiliated with McMaster English Language Development Diploma (MELD) program, the Department of Linguistics and Languages, and the Centre for Advanced Research in Experimental and Applied Linguistics (ARiEAL) at McMaster University. Sadaf Rahmanian is affiliated with the MELD program and ARiEAL at McMaster University. For eye-movement data collection, we are thankful to Joyce Bauman, Laura Beaudin, Cassandra Chapman, Melda Coşkun, Meliha Horzum, Constance Imbault, Alyssa Janes, Richard Mah, Kirandeep Matharu, Anya Minchenkova, Kelly Nisbet, Fareeha Rana, Avneet Sandhu, Theodore Saratsiotis, Bryor Snefjella, and Jennifer Thammasouk. We are grateful to both Victor Kuperman and the Dean's Office, Faculty of Humanities at McMaster University, for critical infrastructure support provided to the Bilingualism Lab in ARiEAL.

Supplementary Material. For supplementary material accompanying this paper, visit https://doi.org/10.1017/S1366728922000542

References

Al-Homoud, F and Schmitt, N (2009) Extensive reading in a challenging environment: A comparison of extensive and intensive reading approaches in Saudi Arabia. *Language Teaching Research* **13**, 383–401.

Ashby, J, Rayner, K and Clifton Jr, C (2005) Eye movements of highly skilled and average readers: Differential effects of frequency and predictability. *The Quarterly Journal of Experimental Psychology Section A* **58**, 1065–1086.

Baayen, RH, Davidson, DJ and Bates, DM (2008) Mixed-effects modeling with crossed random effects for subjects and items. *Journal of Memory and Language* **59**, 390–412.

Baayen, RH and Milin, P (2010) Analyzing reaction times. *International Journal of Psychological Research* **3**, 12–28.

Barnett, AG, van der Pols, JC and Dobson, AJ (2005) Regression to the mean: what it is and how to deal with it. *International journal of epidemiology* **34**, 215–220.

Bates, D, Mächler, M, Bolker, B and Walker, S (2015) Fitting linear mixed-effects models using lme4. *Journal of Statistical Software* **67**, 1–48. doi: 10.18637/jss.v067.i01

Bax, S (2013) The cognitive processing of candidates during reading tests: Evidence from eye-tracking. *Language Testing* **30**, 441–465.

Beglar, D and Hunt, A (2014) Pleasure reading and reading rate gains. *Reading in a Foreign Language* **26**, 29–48.

Beglar, D, Hunt, A and Kite, Y (2012) The effect of pleasure reading on Japanese university EFL learners reading rates. *Language learning* **62**, 665–703.

Benjamini, Y and Hochberg, Y (1995) Controlling the false discovery rate: a practical and powerful approach to multiple testing. *Journal of the Royal statistical society: series B (Methodological)* **57**, 289–300.

Bernhardt, E (2011) *Understanding advanced second-language reading*. Routledge.

Brysbaert, M (2019a) How many participants do we have to include in properly powered experiments? a tutorial of power analysis with reference tables. *Journal of Cognition* **2**, 16. doi: 10.5334/joc.72

Brysbaert, M (2019b) How many words do we read per minute? A review and meta-analysis of reading rate. *Journal of Memory and Language* **109**, 104047.

Brysbaert, M (2021) Power considerations in bilingualism research: Time to step up our game. *Bilingualism: Language and Cognition* **24**, 813–818.

Brysbaert, M and New, B (2009) Moving beyond Kučera and Francis: A critical evaluation of current word frequency norms and the introduction of a new and improved word frequency measure for American English. *Behavior research methods* **41**, 977–990.

Brysbaert, M and Stevens, M (2018) Power analysis and effect size in mixed effects models: A tutorial. *Journal of cognition* **1**, 1–20.

Carver, RP (1990) *Reading rate: A review of research and theory*. Academic Press.

Conklin, K, Alotaibi, S, Pellicer-Sánchez, A and Vilkaitė-Lozdienė, L (2020) What eye-tracking tells us about reading-only and reading-while-listening in a first and second language. *Second Language Research* **36**, 257–276.
Conklin, K and Pellicer-Sánchez, A (2016) Using eye-tracking in applied linguistics and second language research. *Second Language Research* **32**, 453–467.
Conklin, K, Pellicer-Sánchez, A and Carrol, G (2018) *Eye-tracking: A guide for applied linguistics research*. Cambridge University Press.
Cook, AE and Wei, W (2017) Using eye movements to study reading processes: Methodological considerations. In CA Was, FJ Sansoit and BJ Morris (Eds.), *Eye tracking technology applications in educational research*. IGI Global: Hershey, (pp. 27–47).
Cook, AE and Wei, W (2019) What can eye movements tell us about higher level comprehension? *Vision* **3**, 45.
Cop, U, Dirix, N, Drieghe, D and Duyck, W (2017) Presenting GECO: An eyetracking corpus of monolingual and bilingual sentence reading. *Behavior research methods* **49**, 602–615.
Cop, U, Dirix, N, Van Assche, E, Drieghe, D and Duyck, W (2017) Reading a book in one or two languages? An eye movement study of cognate facilitation in L1 and L2 reading. *Bilingualism: Language and Cognition* **20**, 747–769.
Cop, U, Drieghe, D and Duyck, W (2015) Eye movement patterns in natural reading: A comparison of monolingual and bilingual reading of a novel. *PloS one* **10**, e0134008.
Daller, MH and Phelan, D (2013) Predicting international student study success. *Applied Linguistics Review* **4**, 173–193.
Daller, MH and Xue, H (2009) Vocabulary knowledge and academic success: A study of Chinese students in UK higher education. In B Richards, MH Daller, DD Malvern, P Meara, J Milton and J Treffers-Daller (Eds.), *Vocabulary studies in first and second language acquisition: The interface between theory and application*. Springer, (pp. 179–193).
Dirix, N, Vander Beken, H, De Bruyne, E, Brysbaert, M and Duyck, W (2020) Reading text when studying in a second language: An eye-tracking study. *Reading Research Quarterly* **55**, 371–397.
Droop, M and Verhoeven, L (2003) Language proficiency and reading ability in first-and second-language learners. *Reading research quarterly* **38**, 78–103.
Farnia, F and Geva, E (2013) Growth and predictors of change in English language learners' reading comprehension. *Journal of Research in Reading* **36**, 389–421.
Fox, J and Weisberg, S (2019) *An R companion to applied regression* (3rd ed.). Thousand Oaks: CA: Sage.
Frazier, L and Rayner, K (1982) Making and correcting errors during sentence comprehension: Eye movements in the analysis of structurally ambiguous sentences. *Cognitive psychology* **14**, 178–210.
Fuchs, LS, Fuchs, D, Hosp, MK and Jenkins, JR (2001) Oral reading fluency as an indicator of reading competence: A theoretical, empirical, and historical analysis. *Scientific studies of reading* **5**, 239–256.
Geva, E, Xi, Y, Massey-Garrison, A and Mak, JY (2019) Assessing reading in second language learners: Development, validity, and educational considerations. In DA Kilpatrick, RM Joshi and RK Wagner (Eds.), *Reading development and difficulties*. Springer, (pp. 117–155).
Godfroid, A (2020) *Eye tracking in second language acquisition and bilingualism: A research synthesis and methodological guide*. Routledge: New York and London.
Godfroid, A, Ahn, J, Choi, I, Ballard, L, Cui, Y, Johnston, S, Lee, S, Sarkar, A and Yoon, HJ (2018) Incidental vocabulary learning in a natural reading context: An eye-tracking study. *Bilingualism: Language and Cognition* **21**, 563–584.
Gordon, PC, Moore, M, Choi, W, Hoedemaker, RS and Lowder, MW (2020) Individual differences in reading: Separable effects of reading experience and processing skill. *Memory & Cognition* **48**, 553–565.
Grabe, W (2009) *Reading in a second language: Moving from theory to practice*. Ernst Klett Sprachen.
Green, P and MacLeod, CJ (2016) simr: an R package for power analysis of generalised linear mixed models by simulation. *Methods in Ecology and Evolution* **7**, 493–498. Retrieved from https://CRAN.R-project.org/package=simr doi:10.1111/2041-210X.12504
Gulliksen, H (1950) The reliability of speeded tests. *Psychometrika* **15**, 259–269.
Harrington, M and Roche, T (2014) Identifying academically at-risk students in an English-as-a-Lingua-Franca university setting. *Journal of English for Academic Purposes* **15**, 37–47.
Hebert, ME (2017) An examination of reading comprehension and reading rate in university students. *Unpublished doctoral dissertation, University of Alberta, Canada.*
Hirai, A (1999) The relationship between listening and reading rates of Japanese EFL learners. *The modern language journal* **83**, 367–384.
Huffman, J (2014) Reading rate gains during a one-semester extensive reading course. *Reading in a Foreign Language* **26**, 17–33.
Inhoff, A, Weger, U and Radach, R (2005) Sources of information for the programming of short- and long-range regressions during reading. In G Underwood (Ed.), *Cognitive processes in eye guidance*. Oxford University Press: NY, (pp. 33–52).
Jaeger, TF (2008) Categorical data analysis: Away from anovas (transformation or not) and towards logit mixed models. *Journal of Memory and Language* **59**, 434–446.
Jenkins, JR, Fuchs, LS, van den Broek, P, Espin, C and Deno, SL (2003) Sources of individual differences in reading comprehension and reading fluency. *Journal of educational psychology* **95**, 719–729.
Jeon, EH and Yamashita, J (2014) L2 reading comprehension and its correlates: A meta-analysis. *Language learning* **64**, 160–212.
Jiang, X (2011) The role of first language literacy and second language proficiency in second language reading comprehension. *The Reading Matrix* **11**, 177–190.
Joseph, HS and Liversedge, SP (2013) Childrens and adults on-line processing of syntactically ambiguous sentences during reading. *PloS one* **8**, e54141.
Keung, YC and Ho, CSH (2009) Transfer of reading-related cognitive skills in learning to read Chinese (L1) and English (L2) among Chinese elementary school children. *Contemporary Educational Psychology* **34**, 103–112.
Kieffer, MJ (2011) Converging trajectories: Reading growth in language minority learners and their classmates, kindergarten to grade 8. *American Educational Research Journal* **48**, 1187–1225.
Koda, K (2005) *Insights into second language reading: A cross-linguistic approach*. Cambridge University Press.
Koda, K (2008) Impacts of prior literacy experience on second language learning to read. In K Koda and AM Zehler (Eds.), *Learning to read across languages: Cross-linguistic relationships in first- and second-language literacy development*. Routledge New York, NY, (pp. 68–96).
Kuperman, V, Matsuki, K and Van Dyke, JA (2018) Contributions of reader- and text-level characteristics to eye-movement patterns during passage reading. *Journal of Experimental Psychology: Learning, Memory, and Cognition* **44**, 1687–1713.
Kuperman, V, Siegelman, N, Schroeder, S, Alexeeva, A, Acartürk, C, Amenta, S, Bertram, R, Bonandrini, R, Brysbaert, M, Chernova, D. Da Fonseca, SM, Dirix, N, Duyck, W, Fella, A, Frost, R, Gattei, CA, Kalaitzi, A, Lõo, K, Marelli, M, Nisbet, K, Papadopoulos, TC, Protopapas, A, Savo, S, Shalom, DE, Slioussar, N, Stein, R, Sui, L, Taboh, A, Tønnesen, V and Usal, KA (2022) Text reading in English as a second language: Evidence from the multilingual eye-movements corpus (MECO). *Studies in Second Language Acquisition*.
Kuznetsova, A, Brockhoff, PB and Christensen, RHB (2017) lmerTest package: Tests in linear mixed effects models. *Journal of Statistical Software* **82**, 1–26. doi:10.18637/jss.v082.i13
LaBerge, D and Samuels, SJ (1974) Toward a theory of automatic information processing in reading. *Cognitive psychology* **6**, 293–323.
Lao, CY and Krashen, S (2000) The impact of popular literature study on literacy development in EFL: More evidence for the power of reading. *System* **28**, 261–270.
Lervåg, A and Aukrust, VG (2010) Vocabulary knowledge is a critical determinant of the difference in reading comprehension growth between first and second language learners. *Journal of Child Psychology and Psychiatry* **51**, 612–620.
Ling, G, Wolf, MK, Cho, Y and Wang, Y (2014) English-as-a-second-language programs for matriculated students in the United States: An exploratory survey and some issues. *ETS Research Report Series* **2014**, 1–19.
Liversedge, SP, Blythe, HI and Drieghe, D (2012) Beyond isolated word recognition. *Behavioral and Brain Sciences* **35**, 293–294.

McLean, S and Rouault, G (2017) The effectiveness and efficiency of extensive reading at developing reading rates. *System* **70**, 92–106.

Melby-Lervåg, M and Lervåg, A (2014) Reading comprehension and its underlying components in second-language learners: A meta-analysis of studies comparing first-and second-language learners. *Psychological bulletin* **140**, 409–433.

Meseguer, E, Carreiras, M and Clifton, C (2002) Overt reanalysis strategies and eye movements during the reading of mild garden path sentences. *Memory & Cognition* **30**, 551–561.

Muñoz, C and Singleton, D (2011) A critical review of age-related research on L2 ultimate attainment. *Language teaching* **44**, 1–35.

Nakamoto, J, Lindsey, KA and Manis, FR (2007) A longitudinal analysis of English language learners word decoding and reading comprehension. *Reading and Writing* **20**, 691–719.

Nakanishi, T (2015) A meta-analysis of extensive reading research. *TESOL Quarterly* **49**, 6–37.

Nisbet, K, Bertram, R, Erlinghagen, C, Pieczykolan, A and Kuperman, V (2021) Quantifying the difference in reading fluency between L1 and L2 readers of English. *Studies in Second Language Acquisition* **128**. doi: 10.1017/S0272263121000279

Pellicer-Sánchez, A (2016) Incidental L2 vocabulary acquisition from and while reading: An eye-tracking study. *Studies in Second Language Acquisition* **38**, 97–130.

Pellicer-Sánchez, A, Tragant, E, Conklin, K, Rodgers, M, Llanes, A and Serrano, R (2018) L2 reading and reading-while-listening in multimodal learning conditions: An eye-tracking study. *ELT Research Papers* **18**, 1–28.

Pfost, M, Hattie, J, Dörfler, T and Artelt, C (2014) Individual differences in reading development: A review of 25 years of empirical research on Matthew effects in reading. *Review of Educational Research* **84**, 203–244.

Philipp, AM and Huestegge, L (2015) Language switching between sentences in reading: Exogenous and endogenous effects on eye movements and comprehension. *Bilingualism* **18**, 614–625.

Phillips, LM, Norris, SP, Osmond, WC and Maynard, AM (2002) Relative reading achievement: A longitudinal study of 187 children from first through sixth grades. *Journal of Educational Psychology* **94**, 3–13.

Pinheiro, JC and Bates, DM (2000) Linear mixed-effects models: basic concepts and examples. *Mixed-effects models in S and S-Plus*, 3–56.

Priya, K and Wagner, RK (2009) The roles of fluent decoding and vocabulary in the development of reading comprehension. *Beyond decoding: The behavioral and biological foundations of reading comprehension*, 124–139.

Protopapas, A, Sideridis, GD, Mouzaki, A and Simos, PG (2011) Matthew effects in reading comprehension: Myth or reality? *Journal of Learning Disabilities* **44**, 402–420.

Raudszus, H, Segers, E and Verhoeven, L (2021) Patterns and predictors of reading comprehension growth in first and second language readers. *Journal of Research in Reading* **44**, 400–417.

Rayner, K (1998) Eye movements in reading and information processing: 20 years of research. *Psychological bulletin* **124**, 372–422.

Rayner, K (2009) Eye movements in reading: Models and data. *Journal of eye movement research* **2**, 1–10.

Rayner, K, Chace, KH, Slattery, TJ and Ashby, J (2006) Eye movements as reflections of comprehension processes in reading. *Scientific studies of reading* **10**, 241–255.

Rayner, K, Pollatsek, A, Ashby, J and Clifton Jr, C (2012) *Psychology of reading*. Psychology Press.

R Core Team. (2021) *R: A language and environment for statistical computing [Computer software manual]*. Vienna, Austria. Retrieved from https://www.R-project.org/

Redden, E (2013) Conditional admission and pathway programs proliferate. *Inside Higher Ed, January* **3**.

Rizopoulos, D (2006) ltm: An R package for latent variable modelling and item response theory analyses. *Journal of Statistical Software* **17**, 1–25. Retrieved from https://doi.org/10.18637/jss.v017.i05

Robb, TN and Susser, B (1989) Extensive reading vs skills building in an EFL context. *Reading in a Foreign Language* **5**, 239–251.

Roehrig, AD, Petscher, Y, Nettles, SM, Hudson, RF and Torgesen, JK (2008) Accuracy of the DIBELS oral reading fluency measure for predicting third grade reading comprehension outcomes. *Journal of School Psychology* **46**, 343–366.

Sakurai, N (2015) The influence of translation on reading amount, proficiency, and speed in extensive reading. *Reading in a Foreign Language* **27**, 96–112.

Schmidtke, D and Moro, AL (2021) Determinants of word-reading development in English learner university students: A longitudinal eye movement study. *Reading Research Quarterly* **56**, 819–854.

Spichtig, A, Pascoe, J, Ferrara, J and Vorstius, C (2017) A comparison of eye movement measures across reading efficiency quartile groups in elementary, middle, and high school students in the US. *Journal of Eye Movement Research* **10**, 1–17.

Stanovich, KE (2000) *Progress in understanding reading: Scientific foundations and new frontiers*. New York: Guilford.

Stanovich, KE, Nathan, RG and Vala-Rossi, M (1986) Developmental changes in the cognitive correlates of reading ability and the developmental lag hypothesis. *Reading research quarterly*, 267–283.

Suk, N (2017) The effects of extensive reading on reading comprehension, reading rate, and vocabulary acquisition. *Reading research quarterly* **52**, 73–89.

Thorpe, A, Snell, M, Davey-Evans, S and Talman, R (2017) Improving the academic performance of non-native English-speaking students: The contribution of pre-sessional English language programmes. *Higher Education Quarterly* **71**, 5–32.

Trenkic, D and Warmington, M (2019) Language and literacy skills of home and international university students: How different are they, and does it matter? *Bilingualism: Language and Cognition* **22**, 349–365.

Van Viegen, S and Russell, B (2019) More than language–evaluating a Canadian university EAP bridging program. *TESL Canada Journal* **36**, 97–120.

Venables, WN and Ripley, BD (2002) *Modern applied statistics with s* (Fourth ed.). New York: Springer. Retrieved from https://www.stats.ox.ac.uk/pub/MASS4/ (ISBN 0-387-95457-0)

Verhoeven, L, Perfetti, C and Pugh, K (2019) Cross-linguistic perspectives on second language reading. *Journal of Neurolinguistics* **50**, 1–6.

Von der Malsburg, T and Angele, B (2017) False positives and other statistical errors in standard analyses of eye movements in reading. *Journal of Memory and Language* **94**, 119–133.

Vorstius, C, Radach, R, Mayer, MB and Lonigan, CJ (2013) Monitoring local comprehension monitoring in sentence reading. *School Psychology Review* **42**, 191–206.

Wang, M, Perfetti, CA and Liu, Y (2005) Chinese–English biliteracy acquisition: Cross-language and writing system transfer. *Cognition* **97**, 67–88.

Whitford, V and Titone, D (2012) Second-language experience modulates first-and second-language word frequency effects: Evidence from eye movement measures of natural paragraph reading. *Psychonomic Bulletin & Review* **19**, 73–80.

Whitford, V and Titone, D (2017) The effects of word frequency and word predictability during first-and second-language paragraph reading in bilingual older and younger adults. *Psychology and aging* **32**, 158-177.

Wickham, H (2016) *ggplot2: Elegant graphics for data analysis*. Springer-Verlag New York. Retrieved from https://ggplot2.tidyverse.org

Wiederholt, JL and Bryant, BR (2012) *GORT-5 Gray Oral Reading Tests–Fifth Edition*. Austin, TX: PRO-ED.

Xue, J (2021) The developmental trajectory of biliteracy for Chinese–English adult EFL learners: a longitudinal study. *Reading and Writing* **34**, 1089–1114.

Xue, J and Jiang, X (2017) The developmental relationship between bilingual morphological awareness and reading for Chinese EFL adult learners: a longitudinal study. *Reading and Writing* **30**, 417–438.

Yixin, W and Daller, M (2014) Predicting Chinese students academic achievement in the UK. In *Proceedings of the 47th annual meeting of the British association for applied linguistics, learning, working and communicating in a global context* (pp. 217–227).

Zhang, D and Koda, K (2012) Contribution of morphological awareness and lexical inferencing ability to L2 vocabulary knowledge and reading comprehension among advanced EFL learners: Testing direct and indirect effects. *Reading and Writing* **25**, 1195–1216.