DOI: 10.1111/modl.70030

ORIGINAL ARTICLE

The Modern Language Journal

The role of first and second language reading, first language low-level skills, and working memory in second language writing

Judit Kormos[1,2] | **Csilla Bartha**[3]

[1]Department of Linguistics and English Language, Lancaster University, Lancaster, UK

[2]Faculty of Education, University of Ljubljana, Ljubljana, Slovenia

[3]Department of Modern Hungarian Linguistics, Eötvös Loránd University, Budapest, Hungary

Correspondence
Judit Kormos, Department of Linguistics and English Language, Lancaster University, Lancaster, UK.
Email:j.kormos@lancaster.ac.uk

Abstract

Our research investigated how L2 and L1 reading, L1 low-level skills and working memory are related to ratings and the linguistic characteristics (productivity, cohesion, lexical sophistication and diversity, syntactic complexity, and accuracy) of argumentative and narrative texts. The research was conducted in Hungary with 95 secondary school students whose proficiency ranged from pre-intermediate to high upper intermediate. Participants' L1 reading and L1 low-level skills were assessed with validated instruments in Hungarian and working memory with the backward digit span test. We administered three L2 reading comprehension tasks from the Cambridge English First for Schools test. Participants wrote a narrative and an argumentative essay. L2 reading scores were significantly associated with raters' perceptions of writing quality and measures of linguistic performance. The relationship between L1 reading on L2 writing was mediated by L2 reading. L2 reading scores were significantly related to productivity, grammatical accuracy, lexical diversity and sophistication and the organization of the written texts highlighting the substantial role of shared L2 grammatical and lexical resources and importance of reading skills for monitoring and revising L2 written output. Participants with higher L1 reading ability and L1 low-level skills scored higher on spelling and mechanics and organization in the narrative text and used fewer connectives in the argumentative text. The narrative texts of students with lower L1 low-level skills contained more spelling errors. Working memory played a limited role in L2 writing. Implications of findings for researching and teaching L2 reading and writing are discussed.

KEYWORDS
cognition, reading, writing

Second language (L2) reading and writing skills play an important role in the process of language learning as means to gain input and produce output that supports development of L2 skills. They are also key literacy skills that L2 users apply in professional and academic contexts. L2 reading and writing are closely related as they draw on a shared resource of L2 knowledge and skills, cognitive abilities, and language independent background knowledge (Pae, 2019). Reading skills are also involved in producing written texts, particularly at the stage when writers review and revise what they had written so far. There is also a strong mutually supportive, dynamic relationship between first language (L1) and L2 literacy skills (Kim et al., 2021). Yet relatively little research has been conducted on how L1 and L2 reading skills influence L2 writing performance and most existing studies have been carried out in Asian contexts where the participants' L1 has a different orthographic system (e.g., Korean L1 which is alpha-syllabic and L2 English which is alphabetic).

The existing studies have used ratings of essay scores as outcome measures of L2 writing quality and have not considered the finer-grained effects of L1 and L2 reading ability on the linguistic aspects of writing such as syntactic complexity, lexical sophistication and variety, grammatical and spelling accuracy, or the use of cohesive tools.

To our knowledge, no previous research has investigated the additional foundational role of lower-level L1 skills such as word-level decoding and rapid automated naming as well as working memory (WM) in L2 writing performance using both analytic ratings of essay quality and linguistic variables. Studying these effects of L1 and L2 reading and foundational L1 linguistic and cognitive abilities on L2 writing in different orthographic and educational contexts is important because it helps language teachers identify struggling L2 writers and understand how they can assist L2 learners with different literacy and cognitive profiles to improve their writing skills. Studies in this area can also shed light on the potential benefits of L1 literacy instruction and early years' foundational linguistic skill interventions for L2 learning.

Our study aimed to fill these research gaps in a context where the participants' L1 and the target L2 both share an alphabetic writing system. The research was conducted in Hungary with the participation of 95 secondary school students whose proficiency ranged from B1 to low C1 on the Common European Framework of Reference (Council of Europe, 2001). In our research we investigated how L2 and L1 reading, L1 low-level skills, and WM are related to ratings and the linguistic characteristics (productivity, cohesion, lexical sophistication and diversity, syntactic complexity, and accuracy) of argumentative and narrative texts.

REVIEW OF LITERATURE

Relationship between L1 and L2 literacy skills

There are several theoretical frameworks and models that describe the inter-relationship of L1 and L2 literacy skills. Geva and Ryan's (1993) *Common Underlying Processes Framework* suggests that reading development in both monolingual and bilingual children is influenced by a distinct set of individual difference variables. These variables include phonological awareness, verbal WM, central executive functions, rapid automated naming, metacognitive awareness, and background knowledge. This framework also aligns with the *Linguistic Interdependence Hypothesis* (Cummins, 1979), which proposes that L1 and L2 literacy skills are significantly interconnected. Cummins (1979) argued that academic language proficiency in both L1 and L2 draws on a common set of underlying language skills that are interdependent and shared across languages. Therefore, academic language proficiency acquired in L1, including text comprehension and production skills, can be transferred to other languages learned later.

A complementary theory to Common Underlying Processes Framework and the Linguistic Interdependence Hypothesis is the *Linguistic Threshold Hypothesis* that posits that below a certain level of L2 proficiency, readers cannot utilize their L1 reading skills to comprehend L2 texts effectively (e.g., Alderson, 1984; Bernhardt & Kamil, 1995). Grabe and Yamashita's (2022) review of empirical research on the L1–L2 reading relationship concluded that there is sufficient evidence for the Common Underlying Processes Framework and the Linguistic Threshold Hypothesis, but the Linguistic Interdependence Hypothesis can only be supported if L2 proficiency and L1–L2 distance factors (cf. Koda's, 2005 transfer facilitation model) are considered.

In discussions of the L1–L2 reading relationships it is also important to acknowledge that "L2 reading is a complex construct composed of various cognitive, linguistic, conative, and affective variables" (Jeon & Yamashita, 2022, p. 31). On the one hand, as postulated by Geva and Ryan's (1993) Common Underlying Processes Framework, L2 reading is influenced by language-specific versus language general skills and abilities. On the other hand, similar to L1 reading, the development of L2 text comprehension skills is influenced by a variety of affective and socio-cultural factors. In her *Direct and Indirect Effects Model of Reading* (DIER), Kim (2020) demonstrates that there is an interactive relationship between cognitive (e.g., WM and central executive functions), and socio-emotional determinants (socio-emotions or reading affect) of L1 reading comprehension. The results of Jeon and Yamashita's (2022) meta-analysis of L2 reading research also provide evidence for a similar multi-componential model L2 of reading comprehension.

As highlighted by Geva and Ryan's (1993) Common Underlying Processes Framework, lower-level L1 skills such as phonological awareness and rapid automated naming in L1 can significantly contribute to L2 reading ability. However, as literacy skills develop, the predictive role of phonological awareness diminishes (Landerl et al., 2013), and among lower-level L1 skills rapid automated naming skills take over their place as a potentially influential predictor variable (Erdos et al., 2014). As automatized and efficient word-level decoding skills serve as solid foundations for higher-level text comprehension, timed and non-timed real and non-word reading ability in L1 has also been found to be a relevant factor influencing L2 reading comprehension (Alderson et al., 2015; Kormos et al., 2019; Van Gelderen et al., 2004).

The relationship between L1 and L2 writing skills has also been examined from the perspectives of the Linguistic Interdependence and the Linguistic Threshold Hypotheses. In several studies such as in Sasaki and Hirose's (1996) project with Japanese learners, Pae's (2019) and Kim and Pae's (2021, 2023) research with Korean learners of L2 English a significant link between L1 and L2 writing ability was established providing support for the Linguistic Interdependence Hypothesis. However, Carson et al. (1990) found no correlation between L1 and L2 writing for participants with L1 Chinese, while Kim et al.'s (2021) study with Korean learners demonstrated that the L1–L2 writing link is mediated by L2 proficiency. These findings indicate that the successful transfer of L1 writing skills might only be possible above a certain level of L2 competence.

Relationship between reading and writing skills

Processes of reading are inherently involved in producing written texts. Hayes' (2012) cognitive model originally proposed to account for L1 writing processes is one of the most detailed models that can explain how reading and writing processes are interrelated and how they draw on a shared pool of resources. According to the model, writing is composed of three levels: control, resource, and process. The control level encompasses the task schema, which serves as a mental representation of information that the writer utilizes to regulate the interaction between writing processes, motivation, and goals related to planning, writing, and revising. The mental resource consists of attention, working- and long-term memory. This mental resource is a shared one between reading and writing processes because both reading and writing utilize content, metacognitive, linguistic and discourse knowledge, and procedural knowledge for accessing, applying, and integrating these types of knowledge (Fitzgerald & Shanahan, 2000). The process level of writing is further divided into writing processes and the task environment. The task environment includes collaborators and reviewers of the written piece, transcribing technology, digital writing support tools, task materials, written plans, and the text produced thus far. The writing process is managed by the evaluator, which monitors the output of the three other processors: the proposer, the translator, and the transcriber. The evaluator applies reading processes to review, edit, and revise the text produced so far. The proposer is responsible for generating ideas and conceptualizing prelinguistic input for writing. The translator converts these ideas into linguistic form, maintaining this linguistic output in WM so that it is available for the evaluator. If the evaluator deems the language output appropriate, it proceeds to the transcriber for textual transcription. If the language output is found to be inappropriate or inaccurate, the proposer and/or translator is reactivated.

The shared cognitive processes and knowledge sources of reading and writing have also been empirically tested in studies examining the nature of the relationship between reading and writing. As Hayes' (2012) model also proposes, writing and reading processes are interdependent and interactive. Indeed, a considerable number of studies in the field of L1 literacy research suggest that compared to one-directional reading-to-writing or writing-to-reading relationships, models that assume a two-way interactive relationship are better supported by data collected from L1 users (e.g., Berninger & Abbott, 2010). Previous research with L2 learners also indicates that L2 writing and reading skills can mutually support each other. For example, in Lee's (2005) study, Taiwanese university students who engaged in more extensive reading in the L2 wrote better essays. Cho and Butt-Griffler (2015) found that integrated reading and writing instruction had a positive effect on the development of L2 reading and writing skills. In a more recent study with Korean learners of English, Kim and Pae's (2023) regression analyses revealed that L2 reading comprehension predicted L2 writing performance, and L2 writing scores were also significant contributors to L2 text comprehension skills. Kim et al. (2021) tested an interactive L2 writing and reading model using structural equation modeling (SEM) and assumed that this interactive relationship is mediated by the shared knowledge source of L2 vocabulary knowledge. Data collected from undergraduate L2 users of English in Korea confirmed the model and highlighted the important role of vocabulary size in comprehending and producing L2 texts. Further research conducted mostly in Asian contexts and using argumentative writing tasks has also provided evidence for a moderately strong link between L2 reading and writing abilities (e.g., Ito, 2011; Lee, 2020; Zhang, 2021).

The role of L1 reading in L2 writing has only been examined in a small number of studies, and interestingly all these studies have been conducted with Korean L2 learners of English. Both studies by Pae (2019) and Kim and Pae (2023) indicated that L1 reading scores contribute significantly to L2 writing performance, albeit with a small effect size. Kim et al.'s (2022) structural equation model showed that L2 reading ability mediates the effect of L1 reading suggesting that L1 reading skills can only support L2 writing if L2 learners have adequate L2 reading skills.

The role of working memory and L1 low-level skills in L2 writing

WM is a crucial part of cognitive processing in the human mind, responsible for keeping transient memory items active to support the completion of complex cognitive tasks (Barrouillet et al., 2007). Kellogg's (1996) *Componential Model of Working Memory in Writing* has been highly influential in the domains of L1 and L2 writing. Drawing on Baddeley and Hitch's (1974) model of WM, Kellogg posited that writing, as a deliberate and cognitively demanding activity, draws on the central executive, which regulates attention, throughout all writing stages. He suggested that the planning phase might be affected by the capacity limitations of the visuospatial sketchpad, as writers frequently employ visual aids such as diagrams or images for content generation and text structuring. Moreover, transforming ideas into language, along with monitoring and reviewing, could be restricted by capacity constraints of the phonological loop.

Another model addressing the impact of WM on writing was introduced by McCutchen (1996) who argued that WM regulates the efficient interaction among writing processes and that until foundational writing skills, such as handwriting, are sufficiently automatized, higher-order processes—such as planning, translating, and revising—cannot operate concurrently with these lower-level tasks. Berninger and Swanson (1994) expanded this model and hypothesized that novice writers have limited WM resources not only due to the lack of automaticity in handwriting but also due to the cognitive effort required for spelling. This reduces their capacity to focus on other aspects of writing, such as rhetorical structure and the use of syntactically complex and lexically sophisticated language. This line of research contributed to the development of the *Not-so-Simple View of Writing* (Berninger & Winn, 2006), which further examines the role of central executive functions in writing processes and skill development. More recently this model was further refined by Kim and Schatschneider (2017). Their *Direct and Indirect Effects Model of Writing* posits that foundational language skills (such as vocabulary and grammar) and cognitive skills (including WM and attention control) exert both direct and indirect influences on text production. The indirect effects are mediated by higher-order cognitive skills such as inferencing, comprehension monitoring, and perspective taking. SEM of their data indeed confirmed that WM might not have a direct effect on writing quality; instead, its effect is mediated by oral discourse production abilities and transcription skills (handwriting fluency and spelling).

The role of WM in L2 writing has been researched extensively and two recently published reviews give a detailed account of existing findings (Kormos, 2023; Li, 2023). The available empirical studies have yielded mixed results regarding the role of WM in predicting writing quality. (e.g., Kormos & Sáfár, 2008; Michel et al., 2019; Révész et al., 2017; Vasylets & Marín, 2021). A recent systematic review by Li (2023) concluded that WM was not a strong predictor of L2 writing performance measured by holistic ratings (e.g., Kormos & Sáfár, 2008; Michel et al., 2019), but was predictive of specific linguistic dimensions, such as complexity, accuracy, and fluency in L2 writing (e.g., Vasylets & Marín, 2021). For example, Adams and Guillot (2008) and Peng et al. (2021) found a significant link between L2 spelling accuracy and WM capacity. Several studies have also shown that higher WM capacity is associated with a reduced number of grammatical errors (e.g., Bergsleithner, 2010; Mavrou, 2020) and increased syntactic complexity in L2 writing (e.g., Bergsleithner, 2010; Mavrou, 2020; Zabihi, 2018).

However, research findings also reveal that the role of WM capacity might be mediated by L2 proficiency, and individual differences in WM might only limit attentional resources for the accuracy and syntactic complexity of L2 writing at lower levels of proficiency (e.g. Manchón et al., 2023; Vasylets & Marín, 2021). Current research evidence appears to suggest a significant role of WM in lexical sophistication (Révész et al., 2017; Vasylets & Marín, 2021), and a limited influence on lexical diversity (Manchón et al., 2023; Mavrou, 2020; Vasylets & Marín, 2021). Importantly, Li's (2023) systematic review highlights that existing findings on the role of working memory in L2 writing performance might be strongly influenced by the type of WM measures employed. Li provides methodological recommendations to support a more systematic and theoretically grounded selection of assessment tools.

The role of word-level decoding skills and rapid automated naming, which in addition to WM are other relevant factors often predictive of literacy difficulties (for a review see Kormos, 2017), has not been extensively investigated in L2 writing, and most existing studies have been conducted in contexts where children acquire English in the target language context. Results of these studies have shown that pseudoword decoding skills are predictive of L2 English spelling skills (e.g., Chiappe et al., 2002; Wade-Woolley & Siegel, 1997). Jongejan et al.'s (2007) study also demonstrated the significant role of rapid automated naming skills in L2 spelling among bilingual children. In Harrison et al.'s (2016) study with multilingual children in Canada, rapid automated naming skills contributed significantly to ratings of content and text structure, text-level spelling accuracy but not to word-level spelling. In another study, also in the Canadian multilingual context, Herbert et al. (2020) found that L2 English-speaking children who were characterized as "poor decoders," scored significantly below their peers in word-level decoding skills, made higher number of spelling errors, and wrote stories that received lower scores on contextual conventions, language, story composition, and total story writing than those whose word-level decoding skills fell in the typical range. The significant role of word-level decoding skills in spelling and writing in English as an additional language has also been confirmed in a recent study by Booton et al. (2024). However, Sehlström et al. (2022) who examined the L2 English writing skills of L1 Swedish-speaking adolescents in an instructed context did not replicate Herbert et al.'s (2020) results and found that poor L1 word-level decoding skills did not result in lower ratings of L2 text quality.

As shown in the review of literature, relatively few studies have examined how reading abilities in both L1 and L2 impact L2 writing performance, and most of this research has focused on Asian contexts where the learners' first language uses a different orthographic system. Existing research has primarily assessed L2 writing quality using essay scores and has not explored in detail how L1 and L2 reading skills are associated with specific linguistic features of writing such as syntactic complexity, lexical diversity and sophistication, grammatical and spelling accuracy, or cohesive device use. To fill these research gaps, our study addressed the following research questions:

RQ 1. How are L2 and L1 reading, L1 low-level skills, and working memory related to ratings of written argumentative and narrative essays?

RQ2. How are L2 and L1 reading, L1 low-level skills, and working memory related to the linguistic characteristics (productivity, cohesion, lexical sophistication and diversity, syntactic complexity, and accuracy) of written argumentative and narrative essays?

METHOD

Participants

The participants of our study were 95 Hungarian learners of English studying in four different secondary schools in Budapest. The schools were located in different areas of the city resulting in a sample that differed in socio-economic status (outskirts and city center) and were of different type (state-funded, fee-paying private, and church-affiliated schools). The participants' age ranged between 15 and 18 years (M_{age} = 16.1 years). Their level of proficiency was between high B1 to low C1 on the Common European Framework of Reference (Council of Europe, 2001; based on teachers' reports and placement tests). Forty-four participants were male and 51 females. All students spoke Hungarian as their L1, but six students reported that they also used additional languages for communication at home (e.g., Arabic and Chinese). None of the students had spent more than 3 months in an English-speaking country.

Instruments

The L2 writing tasks for this study included a narrative and an argumentative essay. For the narrative essay (150–200 words) participants were asked to compose a story based on six unrelated pictures (see Appendix A) so that the storyline would integrate every picture. In the second argumentative essay, students had to express their disagreement with the statement "Teenagers are too young to teach other people about anything." They were instructed to support their views with relevant examples. The word limit for this task was between 200 and 250 words. Students wrote their essays on a computer with no access to spell checking, dictionary, or any other digital support tools.

To assess L2 reading comprehension, we used a freely downloadable version of the Cambridge First Certificate English for Schools test (B2) (a more detailed description of the L2 and L1 reading instruments can be found in Kormos (2025). We selected this test because it had been designed for a similar age and proficiency group as our participants and the psychometric quality of the test was ensured thorough the test validation processes of Cambridge Assessment UK. The test included three reading tasks and consisted of 22 items. In the first test task, participants read a story and answered six multiple choice questions. The second reading task was an informative text consisting of six paragraphs in which a sentence was missing. Students were asked to decide which sentence out of a choice of seven sentences fits the gap in the given paragraph. In the third reading task four people gave their views on environment protection and participants had to make a choice which person expressed a given opinion (10 items). The internal consistency of the L2 reading test was Cronbach's $\alpha = .802$.

L1 reading comprehension was assessed with the help of two tasks, which were taken from two Hungarian national diagnostic L1 literacy tests (two different years 2014 and 2016) that were designed by the Educational Research Institute of the Hungarian Ministry of Human Resources and Education. We chose the two tasks based on the published reliability and facility values in the national administration and the relevance of the topic for our sample of students. One of the texts was informational and the other was an excerpt from a short story. The item types used to assess text comprehension included multiple choice and short answer items (10 items for the informative and 13 items for the narrative text; Cronbach's $\alpha = .777$).

For measuring L1 low-level skills, we used the 3DM-H software, which is an adaptation of the Dutch digital cognitive test battery 3DM (Blomert & Vaessen, 2009) applied for assessing predictors of L1 reading difficulties (e.g., dyslexia). 3DM-H had been carefully validated and standardized for the Hungarian population and is frequently used by professionals for the official identification of specific learning difficulties in Hungary. From the 11 sub-tests of 3DM-H, we originally intended to apply word-level reading tests (real, rare, and non-word reading) to assess word-level decoding skills, rapid automated naming tests to measure speed of lexical retrieval, and a phoneme deletion test to gain insights into participants' phonological awareness. However, students' performance in the phoneme deletion, real, and rare word reading tests was at a ceiling and did not show sufficient variability in the piloting phase. Consequently, we did not use these tests in the data collection for the main study. To assess word-level decoding ability, we only administered the non-word reading test, in which students had to read a list of pseudowords (i.e., words that conform to the phonotactic regularities of Hungarian but are non-existing words) within the time limit of 30 seconds. In the rapid automated naming task, participants were asked to name a series of (a) letters, (b) numbers, and (c) common objects and animals as fast as possible.

WM capacity was assessed with the backward digit span test, which was administered using the Inquisit Lab software.[1] Although there are limitations in using digit-based tests of WM in L2 writing research (for a discussion see Li, 2023), this test was chosen because in previous studies conducted in the Hungarian context (e.g., Kormos & Sáfár, 2008; Michel et al., 2019), it demonstrated the strongest relationship with L2 writing outcomes. Students were asked to recall a series of digits in reverse order. The length of digits increased or decreased based on the accuracy of recall in previous trials. The total length of the test was 14 trials. Working memory capacity was assessed using the two-error maximum length measure.

Procedures

Ethical approval was obtained from the ethics committee of the first author's university. Written informed consent was provided by the participating students and their guardians prior to the start of the research project. We piloted all instruments and procedures with 30 participants from a secondary school 3 months before the main study. This allowed us to check if the allotted time was sufficient for completing the test tasks and that the tests were neither too difficult nor too easy for the students and to establish which components of the 3DM-H test are meaningful to be administered.

The L2 writing tasks were completed by students in groups (15–20 students per group) in a computer lab of the school. The first author and a trained research assistant were present to supervise that the students do not access any digital support tools. Participants had 50 minutes to write both essays, spending 25 minutes on each of the writing tasks which were administered in a counter-balanced order. After a 15-minute break, students took the backward digit span test also on a school computer. The Inquisit Lab software was used to ensure the standard administration of the test. This session was also supervised by the first author and a trained research assistant. The participants took approximately 10 minutes to finish the test.

The L2 and L1 reading comprehension tests were also completed in groups (15–20 students) but for this phase of the study we used paper versions of the tests. The first author and a trained research assistant supervised the administration of these tests. Students had 35 minutes to finish the L2 reading test, and after a 15 minute break, they completed the L1 reading comprehension test which took 45 minutes.

A trained research assistant administered the non-word reading and rapid automated naming components of the 3DMH test individually to the students. This phase of the study lasted for approximately 15 minutes. Approximately half of the student groups completed the L2 writing tasks and the backward digit span task first, and the L1 and L2 reading comprehensions second with approximately 1 week apart. The other half of the groups took these tests in reverse order with a similar time interval between the sessions. The individual administration of the 3DMH followed these tests approximately 2 weeks later.

Scoring and linguistic analyses

The essays were evaluated based on five criteria: (a) content, (b) organization, (c) language use, (d) vocabulary, and (e) spelling and mechanics. Each criterion was rated using a 4-point scale (see Appendix B). The scoring rubric was adapted from Michel et al. (2019) framework, with modifications to the content and organization criteria to align with the characteristics of narrative and argumentative genres and task requirements. The scale was piloted on a random sample of 10 narrative and 10 argumentative essays by the first author and an expert rater who had PhD in L2 assessment. After independently scoring the essays, they discussed scoring discrepancies and interpretation issues at various scale points. Minor modifications to the rating scale were then made, and following this, the expert rater trained a research assistant with an undergraduate degree in linguistics and modern foreign languages on the same randomly selected dataset. After the training, the expert rater and the trained research assistant scored all the essays. Inter-rater reliability calculated using Spearman rank-order correlations exceeded the required agreement threshold of .60 for all sub-criteria and final scores for both types of written tasks (see Table C1 in Appendix C).

A trained research assistant scored the L1 and L2 reading tests using the official answer keys for the tests. The first author checked a random sample of 10 tests. Total test scores were calculated based on the number of correctly answered items. The trained research assistant evaluated the non-word and rapid automated naming tests using the 3DMH interface and the software manual. The score for L1 non-word reading ability was the number of correctly read pseudowords within 30 seconds and the times needed to name digits, letters, and objects served as the measures of naming speed. Backward digit span was calculated automatically by the Inquisit lab software.

Following our previous analysis of this dataset (see Kormos & Suzuki, 2024), we used principal component analysis (PCA) to create factor scores for the linguistic measures (see details below). This analysis revealed that the four productivity measures—number of content words, function words, clauses and T-units—formed a single factor. This factor had an

TABLE 1 Description of task performance measures used in the study.

Linguistic aspect	Measure	Tool
Productivity (one factor)	Number of content words	TAALED
	Number of function words	
	Number of clauses	
	Number of T-units	
Cohesion: Semantic overlap (one factor)	Synonym overlap sentences (nouns)	TAACO
	Latent semantic analysis (sentence)	
	Word2vec similarity (sentence)	
Cohesion:	Frequency of all connectives	TAACO
Lexical diversity (one factor)	hdd42 all word types	TAALED
	hdd42 function words	
	MTLD function words	
Lexical sophistication: Frequency (one factor)	COCA Academic Frequency Log Function Words	TAALES
	COCA Academic Frequency Log Content Words	
	COCA Academic Bigram Frequency Log	
Lexical sophistication: Lexical retrieval speed (one factor)	Lexical decision time Z score content words	TAALES
	Lexical decision time SD content words	
	Word naming response time Z score content words	
	Word naming response time SD score content words	
Lexical complexity	MRC familiarity—content words	TAASSC
Syntactic complexity: Clausal complexity	Mean length of T-unit	TAASSC
Syntactic complexity: Syntactic dependency	Dependents per clause	TAASSC
	Adverbial modifiers per clause	
	Average construction frequency—academic (log transformed)	
Accuracy	Frequency of grammatical errors	GPT-4o
	Frequency of spelling errors	
	Frequency of punctuation errors	

Abbreviations: TAACO, Tool for the Automatic Analysis of Cohesion; TAALED, Tool for the Automatic Analysis of Lexical Diversity; TAALES, Tool for the Automatic Analysis of Lexical Sophistication; TAASSC, Tool for the Automatic Analysis of Syntactic Sophistication and Complexity The tools are freely available on the website "NLP for the Social Sciences" (https://www.linguisticanalysistools.org/).

eigenvalue of 3.13, explaining 78.34% of variance in the argumentative task (KMO = .737, Bartlett's Test $p < .001$), and an eigenvalue of 3.28, explaining 82.55% of variance in the narrative task (KMO = .765, Bartlett's Test $p < .001$) (see Table 1 for a list of measures).

Cohesion metrics were further divided into semantic overlap features and the overall frequency of connectives. Semantic overlap measures, which included Overlap Sentences (nouns), Latent Semantic Analysis (sentence), and Word2vec Similarity (sentence), formed a single factor with an eigenvalue of 1.87 (62.47% variance) for argumentative tasks (KMO = .652, Bartlett's Test $p < .001$) and an eigenvalue of 1.63 (53.96% variance) for narrative tasks (KMO = .605, Bartlett's Test $p < .001$). The overall frequency of connectives was used as a single additional variable.

Lexical diversity was evaluated using the hdd42-index for all words, function words, and MTLTD for function words, together forming a factor with an eigenvalue of 2.25 (75.02% variance) in argumentative tasks (KMO = .725, Bartlett's Test $p < .001$), and an eigenvalue of 2.26 (75.43% variance) in narrative tasks (KMO = .716, Bartlett's Test $p < .001$). Lexical sophistication indices were split into two factors: a lexical frequency factor (including COCA Academic Frequency Log for function and content words, and COCA Academic Bigram Frequency Log) and a lexical retrieval speed factor (Lexical decision time Z score, Lexical decision time SD for content words, Word naming response time Z score, and Word naming response time SD for content words). In the argumentative task, their respective eigenvalues and variances were 2.08 (69.38% KMO = .585, Bartlett's Test $p < .001$) and 2.59 (64.82% variance, KMO = .675, Bartlett's Test $p < .001$); in the narrative task, they were 1.94 (64.78% variance, KMO = .597, Bartlett's Test $p < .001$) and 2.08 (69.38% variance, 1.94 (64.78% variance, KMO = .651, Bartlett's Test $p < .001$). Lexical complexity was operationalized by the single variable of MRC familiarity for content words, which was based on familiarity ratings in the MRC database (Coltheart, 1981).

Syntactic complexity was measured at the clausal level by the mean length of T-units as well as with a syntactic dependency factor comprising dependents per clause, adverbial modifiers per clause, and average academic construction frequency (log transformed). The syntactic dependency factor had an Eigenvalue of 1.85 (61.72% variance, KMO = .662, Bartlett's Test $p < .001$) for argumentative tasks and 1.76 (58.90% variance, KMO = .589, Bartlett's Test $p < .001$) for narrative tasks.

Accuracy was assessed by calculating the number of grammatical, spelling, and punctuation errors per 100 words using the GPT-4o model via the OpenAI API. We used the Google Collab programming interface in Python to process the text of the essays following the procedures and prompts reported by Mizumoto et al. (2024). Twenty randomly selected essays (10 narrative and 10 argumentative) were manually coded by the first author for the number grammar and spelling errors. The Spearman rank-order correlation between the Chat GPT-4o and manual human coding was $\rho = 0.84$ suggesting that the Chat GPT-4o ratings were reliable.

Statistical analyses

To answer our first research question how L2 and L1 reading, L1 low-level skills, and WM are related to ratings of written argumentative and narrative essays, we first conducted correlation analyses among these variables. We then used SEM to examine these relationships. Our original intention was to include all four predictors of L2 writing scores into an SEM model, but because correlations of L1 low-level skills and WM to L1 and L2 reading were non-significant, and the original more complex SEM model failed to converge, we only modeled the relationship of L2 writing scores and L2 and L1 reading using SEM. We first estimated the latent variables of L2 writing scores for the two tasks using the raw analytic rating scores awarded by the two raters with the help of confirmatory factor analysis (Table C2 in Appendix C shows the fit values for the latent variables of the study). Based on our previous analysis of the dataset (Kormos et al., 2024), we assumed that the L2 writing scores were unidimensional and formed one factor for both writing tasks. The latent variable of L2 reading was estimated using the scores for the three reading tasks and the latent variable of L1 low-level skills based on the performance on the non-word reading, and the rapid automated naming speed of letters, numbers, and pictures. As there were only two tasks in the L1 reading test, we did not conduct confirmatory factor analysis for the latent variable of L1 reading. As the L2 writing scores and L1 and L2 reading scores were not normally distributed (for descriptive statistics, see Tables C3 and C4 in Appendix C), we used robust maximum likelihood estimation (Hu & Bentlar, 1998) for building the SEM models. Based on Hu and Bentler (1998), we used the following goodness of fit indices: comparative fit index (CFI $\geq .90$), Tucker–Lewis index (TLI $\geq .90$), and root mean square error of approximation (RMSEA $\leq .08$).The model was estimated through the *cfa* function in the *lavaan* package (Rosseel, 2012), using R statistical software 4.0.2 (R Development Core Team, 2020).

To answer our second research question, we applied factor analyses to merge linguistic performance metrics whenever feasible (see Table C5 in the Appendix C for factor loadings). To reduce the risk of Type I error resulting from performing numerous statistical tests, PCA was utilized to determine if

TABLE 2 Descriptive statistics for the measures of L2 and L1 reading, working memory, and L1 low-level skills.

	Min	Max	Mean	SD
L2 reading total score	4	22	11.322	4.413
L1 reading score	2	22	14.945	4.234
Backward digit	3	8	5.396	1.333
Low-level L1 factor score	−2.660	2.653	0.000	1.00

TABLE 3 Descriptive statistics for the writing scores.

Variable	Task	Min	Max	Mean	SD
Content	ARG	1	4	2.627	0.051
	NAR	1	4	2.865	0.057
Organization	ARG	1	4	2.614	0.051
	NAR	1	4	2.635	0.056
Language use	ARG	1	4	2.785	0.049
	NAR	1	4	2.739	0.048
Vocabulary	ARG	1	4	2.895	0.051
	NAR	1	4	3.078	0.045
Spelling and mechanics	ARG	1	4	2.825	0.047
	NAR	1	4	2.739	0.043

variables converged on a common factor. When such a factor was identified, a regression factor score was generated. Correlation analyses were employed to explore the relationship between linguistic task performance measures and L2 and L1 reading, L1 low-level skills, and WM. Effect sizes were interpreted based on Plonsky and Oswald's (2014) guidelines (small effect size: $r \approx 0.25$, medium effect size: $r \approx 0.40$; large effect size: $r \approx 0.60$).

RESULTS

Preliminary analyses

Prior to answering the research questions, we report the descriptive statistics for all our variables and the intercorrelations of L1 and L2 reading tests, and the measures of low L1 skills and WM. As can be seen in Tables 2 and 3, there was relatively large variability in all the variables of the study and mean values of students' performance in the L1, L2 reading and writing tests were around the middle point between the minimum and maximum possible scores. The mean values for the ratings on the narrative and argumentative tasks were similar across all five criteria.

We observed only a few significant intercorrelations between L1 and L2 reading, L1 low-level skills, and the backward digit span test (see Table 4). L2 and L1 reading were significantly correlated with a small effect size, and a similar significant correlation between L1 low-level skills and backward digit span also with a small effect size was found. Neither backward digit span nor L1 low-level skills factor scores were significantly related to L1 or L2 reading test scores.

TABLE 4 Intercorrelations of measures of L2 and L1 reading, working memory, and L1 low-level skills.

	L1 reading		Backward digit span		L1 low-level skills	
	r	*p*	*r*	*p*	*r*	*p*
L2 reading	.279	.004	.102	.264	.149	.166
L1 reading			.130	.189	.182	.079
Backward digit span					.246	.028

TABLE 5 Correlations of L2 writing scores with measures of L2 and L1 reading, working memory, and L1 low-level skills.

		L2 reading		L1 reading		Backward digit span		L1 low-level skills	
Scale	**Task**	*r*	*p*	*r*	*p*	*r*	*p*	*r*	*p*
Content	ARG	.267	.010	.158	.129	.026	.807	.053	619
	NAR	.197	.058	.113	.279	.157	.138	.138	.191
Organization	ARG	.362	<.001	.177	.090	.025	.813	.110	.299
	NAR	.382	<.001	.206	.048	.172	.102	.311	.003
Vocabulary	ARG	.517	<.001	.149	.154	.112	292	.111	.297
	NAR	.408	<.001	.176	.091	.154	.146	.067	.526
Language use	ARG	445	<.001	.052	.618	.073	.489	.179	.080
	NAR	.475	<.001	.113	.279	.030	.780	.012	.907
Spelling and mechanics	ARG	.183	.080	.133	.202	.114	.283	.171	.106
	NAR	.335	<.001	.222	.033	.126	.235	.296	.004

The relationship between essay ratings and L2 and L1 reading, L1 low-level skills, and working memory

Our first research question asked how L2 and L1 reading, L1 low-level skills, and WM were related to ratings of written argumentative and narrative essays. The correlational analysis revealed that L2 reading scores were significantly correlated with ratings of organization, vocabulary, and language use in both tasks with effect sizes ranging from small to moderate (see Table 5). Ratings for the content of the argumentative task and spelling and mechanics in the narrative task were also significantly related to L2 reading performance with a small effect size, but no significant correlations between content ratings for the narrative task and spelling in the argumentative task were found. L1 reading scores and L1 low-level scores were significantly correlated with raters' evaluation of organization and spelling and mechanics in the narrative task. No other significant correlations emerged between essay scores and L1 low-level skills, and no significant relationships between WM and essay ratings were found.

Our original aim was to test the contribution of L2 and L1 reading, L1 low-level skills, and WM to L2 essay ratings in a hypothetical SEM model. As no significant relationship between backward digit span scores and L2 ratings emerged and because the SEM model including the latent variables of WM and L1 low-level skills failed to converge, we only estimated an SEM model that included L1 and L2 reading as predictors of L2 writing scores. Based on previous studies (e.g., Kim et al., 2021; Pae, 2019) and theoretical considerations (Kim & Schatschneider, 2017; Schoonen, 2022), our theoretical model was a path model that hypothesized that L1 reading had positive effect on L2 reading which in turn was positively associated to the latent variable of L2 writing (see Figure 1) for the hypothetical model. Based on Kim and Pae's (2023) study we also tested an alternative hypothetical model in which L1 reading was assumed to have a direct as well as indirect effect on L2 writing (see Figure 2).

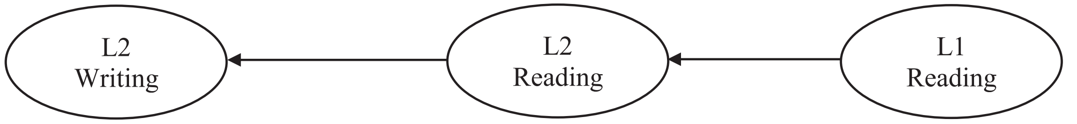


FIGURE 1 The hypothetical model of the contribution of L1 and L2 reading to L2 writing.

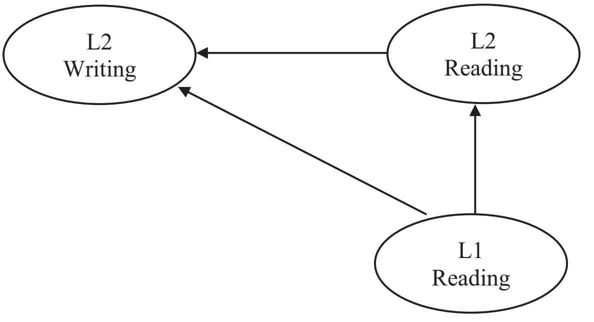


FIGURE 2 The alternative hypothetical model of the contribution of L1 and L2 reading to L2 writing.

TABLE 6 Model fit indices for final and alternative structural equation models for the argumentative and narrative tasks.

Model	Task	df	χ^2	*p*	CFI	TLI	SRMR	RMSEA [90% CI]
Final	Argumentative	45	327.291	<.001	0.924	0.901	0.057	0.068 [0.036, 0.095]
	Narrative	45	338.291	.004	0.924	0.917	0.057	0.078 [0.049, 0.121]
Alternative hypothetical	Argumentative	45	327.291	<.001	0.927	0.897	0.056	0.092 [0.050, 0.132]
	Narrative	45	338.807	.010	0.922	0.891	0.059	0.090 [0.047, 0.129]

The model fit indices for the hypothesized models are summarized in Table 6, and the final models for the two tasks are also visually presented with standardized regression coefficients in Figure 3. No residual covariance between observed variables was assumed in the models. The fit indices for the alternative hypothetical model that assumed both a direct and an indirect effect of L1 reading on L2 writing were below the values of the final model (see Table 6) and indicated a non-significant direct link between L1 reading and argumentative scores ($\beta = -.219$, $p = .353$, 95% CI [−0.680, 0.242]) and narrative essay scores ($\beta = .015$, $p = .943$, 95%CI [0.404, 0.434]). Therefore, this model was rejected. The indices for the final model indicated that the models for both tasks were appropriate fit for the data and hence these measurements models were considered as the final models for our study. In the model for the argumentative task, 44.1% of the variance, and in the narrative task, 34.3% of the variance was explained by the predictors. The final SEM models indicated that in both tasks, L1 reading predicted L2 reading positively (argumentative task: $\beta = .450$, $p = .008$, 95%CI [0.134,0.874]; narrative task: $\beta = .483$, $p = .003$, 95%CI [0.186, 0.920]), which in turn had a significant positive relationship with L2 writing (narrative task: $\beta = .663$, $p < .001$, 95%CI [0.359, 0.906]), argumentative task: $\beta = .594$, $p < .008$, 95%CI [.159, .887]).

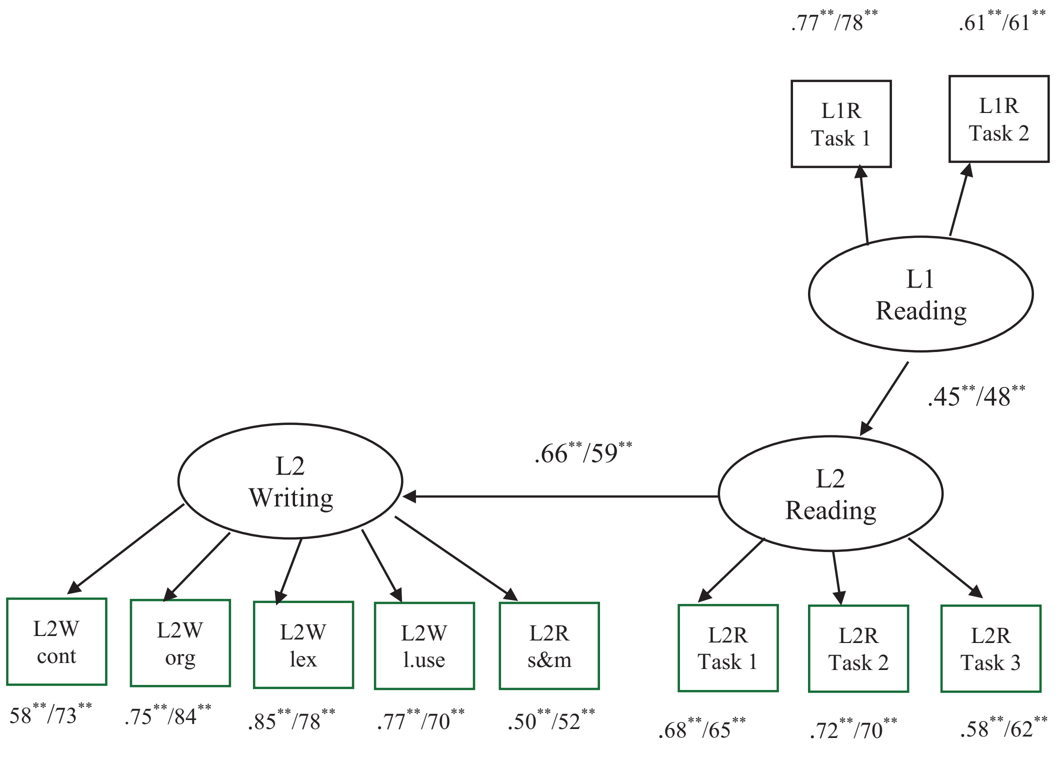


FIGURE 3 The final model of the contribution of L1 and L2 reading to L2 writing in the argumentative and narrative task. Values for the argumentative task are presented first, and values for the narrative task second. Circles represent latent variables. Squares represent observed variables. $^{**}p < .001$. [Color figure can be viewed at wileyonlinelibrary.com]

The relationship between the linguistic features of essays and L2 and L1 reading, L1 low-level skills, and working memory

Our second research question asked how L2 and L1 reading, L1 low-level skills, and working memory were related to the linguistic characteristics (productivity, cohesion, lexical sophistication and diversity, syntactic complexity, and accuracy) of written argumentative and narrative essays. Except for the backward digit span score, all other variables (L2 reading, L1 reading, and L1 low-level skills) were significantly and positively correlated with the productivity factor score in the argumentative task albeit with a small effect size (Table 7). In the narrative task, only L2 reading showed a significant positive correlation to L2 writers' productivity also with a small effect size (see Table 8). As regards the cohesion measures, the semantic overlap factor score was not significantly associated with L2 and L1 reading, L1 low-level skills, or WM in either of the tasks. Participants with higher L1 reading scores and L1 low-level skills used fewer connectives in the argumentative task (small effect size), whereas in the narrative task those who scored higher in L2 reading used fewer connectives (small effect size in both tasks). In terms of lexical measures, higher L2 reading scores were associated with higher lexical diversity, the use of words that take longer to retrieve from memory and content words that are lower in MRC familiarity value in both tasks. The effect size values for these correlations were close to or above moderate effect size. In the narrative task, students with higher L2 reading performance used less frequent words in their stories (small effect size). No significant links between any of the syntactic measures and L2 and L1 reading, L1 low-level skills, or WM were observed in either of the tasks. The narrative and argumentative essays of participants who had higher L2 reading scores contained significantly fewer grammatical errors (close to or above moderate effect size threshold). Students who had stronger L1 low-level skills made fewer spelling mistakes in the narrative task.

TABLE 7 Correlations of linguistic measures with L2 and L1 reading, working memory, and L1 low-level skills in the argumentative task.

Measure	L2 reading		L1 reading		Backward digit span		L1 low-level skills	
	r	*p*	*r*	*p*	*r*	*p*	*r*	*p*
Productivity factor	.267	.006	.269	.005	.148	.134	.222	.032
Cohesion measures								
Semantic overlap factor	.165	.091	.113	.248	.003	.979	.133	.202
Frequency of all connectives	−.176	.072	−.272	.005	−.029	.767	−.217	.036
Lexical measures								
Lexical diversity factor	.420	<.001	−.047	.634	.142	.150	.104	.320
Lexical frequency factor	−.035	.722	.065	.511	−.076	.441	−.149	.152
Lexical retrieval speed factor	.395	<.001	.016	.874	.071	.472	−.087	.404
MRC familiarity—content words	−.414	<.001	−.072	.464	−.082	.409	.070	.502
Syntactic measures								
Mean length of T-unit	.021	.832	−.006	.954	−.005	.958	.022	.833
Syntactic dependency factor	.090	.360	−.056	.566	.008	.939	.019	.858
Accuracy measures								
Frequency of grammatical errors	−.504	<.001	−.082	.462	−.082	.465	−.134	.260
Frequency of spelling errors	−.186	.093	−.190	.086	−.263	.018	−.217	.065

Note: Two-tailed tests.

TABLE 8 Correlations of linguistic measures with L2 and L1 reading, working memory, and L1 low-level skills in the narrative task.

Measure	L2 reading		L1 reading		Backward digit span		L1 low-level skills	
	r	*p*	*R*	*p*	*r*	*p*	*r*	*p*
Productivity factor	.313	.002	.092	.380	.146	.168	.067	.550
Cohesion measures								
Semantic overlap factor	−.114	.275	−.017	.870	−.146	.166	−.074	.507
Frequency of all connectives	−.266	.010	.032	.764	.097	.359	.019	.868
Lexical measures								
Lexical diversity factor	.398	<.001	.034	.747	.155	.141	−.056	.614
Lexical frequency factor	−.222	.032	−.064	.543	−.210	.046	−.016	.883
Lexical retrieval speed factor	.494	<.001	.077	.463	.124	.241	−.001	.996
MRC familiarity—content words	−.381	<.001	-.026	.805	−.086	.416	−.050	.654
Syntactic measures								
Mean length of T-unit	−.029	.782	.057	.585	.065	.541	−.029	.796
Syntactic dependency factor	.174	.094	.003	.980	.041	.697	−.023	.837
Accuracy measures								
Frequency of grammatical errors	−.389	<.001	−.069	.572	−.081	.514	−.044	.736
Frequency of spelling errors	−.136	.266	−.153	.209	−.143	.247	−.265	.039

Note: Two-tailed tests.

DISCUSSION

Our study examined how L2 and L1 reading, L1 low-level skills, and WM are related to scores awarded by raters (RQ1) and the linguistic characteristics (productivity, cohesion, lexical sophistication and diversity, syntactic complexity, and accuracy) of written argumentative and narrative essays (RQ2). Overall, the results of SEM and the correlational analyses show a significant relationship between L1 and L2 reading performance and lend support for both the Common Underlying Processes Framework (Geva & Ryan, 1993) and the Linguistic Interdependence Hypothesis (Cummins, 1979). These theoretical frameworks highlight that L1 text comprehension skills serve as an important foundation for L2 reading skills (see also Sparks & Ganschow's (1993), Linguistic Coding Differences Hypothesis) and that L2 users can transfer higher-order text comprehension skills such as inferencing, identifying main ideas, and focusing selectively (Grabe & Yamashita, 2022) from their L1. Although the correlational analysis suggests that the L1–L2 reading relationship has a small effect size, the latent variable approach used in the SEM analysis shows that around 20% of the variance in L2 reading is explained by L1 text comprehension. Given that a multitude of other variables, including L2 proficiency, background knowledge, literacy experiences, affective and socio-economic factors can account for L2 reading outcomes (cf. Grabe & Yamashita, 2022; Jeon & Yamashita 2022), the 20% explained variance can be viewed as considerable, particularly because L1 reading comprehension skills can be improved through systematic and targeted L1 literacy instruction. Our findings also demonstrate that because word- and sentence-level text decoding skills might be automatized in the investigated age group of L2 learners, L1 low-level skills and WM resources seem to play a limited role in L1 and L2 text comprehension (for a more detailed discussion see Kormos, 2025).

Overall writing scores, content, and productivity

The SEM results of our study showed that L2 and L1 reading jointly explained a relatively large proportion of variance in L2 writing scores (argumentative task: 44.1%, narrative task: 34.3%) and that L1 reading ability had a significant effect on L2 writing achievement with the mediation of L2 text comprehension. The significant link observed between L2 reading performance and L2 essay ratings is in line with the results of studies by Ito (2011), Kim et al. (2022), Lee (2020), Pae (2019), and Zhang (2021) with argumentative tasks, and findings of Pae's (2018) research with narrative tasks. The non-significant direct relationship between L1 reading and L2 writing stands in contrast to the results of Kim and Pae's (2023) study conducted with Korean EFL learners. However, they estimated the existence of this relationship using regression analysis which, compared to SEM, is a less robust approach and does not account for measurement errors.

From a cognitive processing perspective, these findings demonstrate that during reading and writing L2 learners rely on a shared base of domain-related knowledge (e.g., vocabulary), meta-knowledge (e.g., functions and purposes of reading/writing), procedural knowledge of accessing and applying these knowledge sources (cf. Fitzgerald & Shanahan's, 2000, Shared Cognition Model) and higher-order cognitive skills (cf. Kim & Schatschneider's, 2017, Direct and Indirect Model of Writing). Higher levels of L2 reading skills also entail the ability to identify key information and form mental representations of the text, which can then be applied effectively for selecting relevant content and creating a conceptual plan for written compositions (Kim et al., 2021). The SEM models in which the effect of L1 reading ability on L2 writing is moderated by L2 reading also provide indirect support for the Linguistic Threshold Hypothesis (Alderson, 1984) and suggest the transfer of L1 literacy skills can only be effective if L2 learners have acquired a certain level of L2 proficiency.

The overall effect of WM seems to be small in the L2 texts produced by our participants. Although theoretical models such as Kellogg's (1996), Kim and Schatschneider's (2017) Direct and Indirect Model of L1 Writing, and Kormos' (2023) Task-Mediated Cognitive Model of L2 Writing and Writing to Learn would predict that WM capacity and attention regulation abilities exert an important influence on the cognitive processes of L2 writing, no significant relationship between WM capacity and writing scores was detected in our research. These findings are similar to a large number of studies that have examined the relationship between raters' evaluation of L2 writing and WM (e.g., Kim et al., 2021; Michel et al., 2019; Vasylets & Marín, 2021). On the one hand, as argued by Kormos (2023), several task and learner internal variables such as L2 proficiency might mediate the impact of WM on L2 writing processes and the written product. Indeed, previous research that found significant links between WM and L2 writing scores was conducted with lower-level L2 learners (e.g., Li et al., 2024; Mujtaba et al., 2021). On the other hand, Li's (2023) systematic review demonstrated that WM effects were mostly observable in the linguistic characteristics of texts produced by L2 writers.

It is interesting to note that the explanatory power of L1 and L2 reading is smaller in the narrative essay ratings than in those of the argumentative essays. One of the reasons for this difference might be that the argumentative essay required that participants express their own arguments in the L2 and come up with relevant support, whereas in the narrative task the storyline was partly determined by the visual prompts, but students could also adjust the content to their existing L2 linguistic resources.[2] Therefore, the shared L2 knowledge base between L2 reading and writing, such as vocabulary and grammar knowledge, might have played a smaller role in composing the narrative than the argumentative text. It is also possible that a higher level of automaticity in recalling and applying L2 knowledge, which is also necessary for skilled L2 reading, might have freed up attentional resources for content planning in the cognitively more demanding argumentative task. This might have resulted in a stronger association between content scores and L2 reading in the argumentative task, as opposed to the non-significant link of L2 reading to raters' evaluations of content in the narrative task.

The number of units of language produced within a given timeframe has been found to be a strong predictor of L2 writing performance in previous studies (e.g., Kojima and Kaneta's [2022] meta-analysis). Although a range of words was given as a required length for the essays, there was still a considerable variability in the productivity in the written output of our participants. In our dataset, we also found a significant relationship between L2 reading and productivity in both written tasks. The shared variance might be explained by the level of automaticity and proceduralization of L2 knowledge that is required for both fluent and accurate L2 text comprehension and the effortless linguistic encoding of relevant content in L2 writing. In the argumentative task, higher performance on the L1 reading and L1 low-level skills test was also associated with more productivity in writing. It is possible that the generation of arguments and supporting evidence for this task might have been facilitated by more extensive L1 reading experience and higher-level background knowledge gained through L1 reading. From the correlational analysis it is not apparent whether the effect of L1 low-level skills on L2 productivity is a direct or indirect one. However, we would hypothesize that better L1 low-level skills such as word-level decoding and more efficient lexical retrieval might result in the ability to write longer argumentative texts in the L2 with the mediation of L1 and L2 reading skills. As the content of the narrative was partly predetermined by the visuals provided, L1 low-level skills, L1 reading ability, and experience might have played a less important role in the productivity of writing in this task.

Vocabulary and language use scores, accuracy, and lexical and syntactic measures

The examination of the inter-correlations of L1 and L2 reading scores with the analytic ratings of essay quality and linguistic characteristics of the participants' texts yields further insights into areas where L1 and L2 text comprehension and writing might draw on a shared knowledge and skills, and L1 and L2 reading experience might support composing L2 texts. As regards analytic essay ratings, we already explained that except for content scores in the narrative task and spelling and mechanics scores in the argumentative task, all analytic rating criteria were significantly associated with L2 reading scores. The effect sizes for significant correlations indicate that the strongest effects of L2 reading comprehension can be observed in raters' evaluations of vocabulary and language use in both essay types. The analysis of the linguistic features of the compositions also shows significant correlations between L2 reading scores and lexical diversity and the use of words that take longer to retrieve and have lower familiarity index. This suggests that the depth and breadth of L2 vocabulary knowledge is a shared knowledge base that underlies effective L2 text comprehension and the accurate and varied use of sophisticated L2 vocabulary in writing (Kim et al., 2021).

The only significant relationship between WM capacity and L2 writing measures was found between the lexical frequency factor score and backward digit span in the narrative task. The narrative texts produced by participants with higher backward digit span score contained words that were less frequent. Although there is limited previous research on the role of WM in lexical aspects of L2 writing, some existing studies have found similar results (e.g., Révész et al., 2017; Vasylets & Marín, 2021) as our study. It is important to note, however, that a significant effect of WM was only detected in the narrative task which, based on Kormos' (2023) Task-Mediated Cognitive Model of L2 Writing and Writing to Learn, suggests that the attentional and cognitive demands of the task might mediate the impact of WM in lexical selection processes during L2 writing.

As regards language use, the finding that better L2 readers received higher scores on language use and made fewer grammatical errors in their narrative and argumentative tasks demonstrates that grammatical knowledge is key to decoding L2 texts at the sentence level as well for grammatical encoding in the linguistic formulation stage of L2 writing (cf. Lee, 2020). Reading ability is an important resource in models of L1 writing (cf. Hayes (2012), and similarly, higher levels of L2 reading ability might also assist L2 writers to monitor, edit, and revise their texts more effectively.

Although there was a significant link between L2 reading and language use scores in both tasks, no relationship between measures of syntactic complexity and L2 reading comprehension was detected in our dataset. On the one hand, unlike human raters, the analysis of syntactic complexity does not take into account whether L2 learners use these grammatical constructions accurately, which is also reflected in the weak correlations between writing scores and syntactic complexity measures reported in Kojima and Kaneta's (2022) meta-analysis. On the other hand, task and genre characteristics might require the use of specific syntactic constructions (Yoon & Polio, 2017) which might result in less variability in the type and complexity of grammatical structures in a given task.

Spelling and mechanics

In the narrative task, L2 and L1 reading and L1 low-level skills were significantly correlated with raters' evaluations of spelling and mechanics, and participants with better L1 low-level skills were found to make fewer spelling errors in their texts. In the argumentative task, however, the only significant correlation observed was between backward digit span and the frequency of spelling errors. Several explanations for these task-related differences might be possible. On the one hand, in the argumentative task it might have been easier for L2 learners to avoid using words they were not sure how to spell and adjust their texts to their orthographic knowledge. On the other hand, as the content-related cognitive demands might have been lower in the narrative task, the participants might have had more attentional resources available for monitoring the accuracy of their spelling. This might have resulted in more variability in spelling performance based on learner–internal factors. Under the pressures of the higher cognitive demands of the argumentative task, WM resources might have played an important role in orthographic processing, which is in line with findings in earlier studies in the field of L1 (e.g., Ahmed et al., 2022) and L2 writing (e.g., Adams & Guillot, 2008; Peng et al., 2021). Significant links between L1 reading and L1 low-level skills and raters' evaluation of L2 spelling and mechanics in the narrative task also lend support to a number of previous studies that have found that L2 learners and bilingual children with L1 literacy-related difficulties tend to find orthographic processing challenging when producing L2 texts (e.g., Fazio et al., 2021; Helland & Kaasa, 2005;

Łockiewicz & Jaskulska, 2016; Palladino et al., 2016; van Setten et al., 2017). These results highlight that L2 spelling skills draw on a well-established and stable representations of the orthographic form of words in the mental lexicon and foundational language skills such as word-level decoding and efficient lexical access are key to accurate spelling in the L2.

Ratings of organization and cohesion measures

L2 reading skills significantly predicted raters' scores of organization in both tasks. However, higher performance on the L1 reading and L1 low-level skills test was associated with higher ratings of organization in the narrative tasks, but not in the argumentative task. In terms of organization, it is possible that participants who were more skilled readers in their L1 and were likely to read more L1 texts in the narrative genre (e.g., fiction), were more familiar with the organizational genre characteristics of narratives and hence were able to formulate better organized and more coherent stories in their L2. Interestingly, the significant effect of L1 on L2 narrative text organization could only be observed in the evaluations of human raters, but not in the linguistic features that express cohesive links. Therefore, it is possible that knowledge of genre characteristics and ability to form a discourse model of a narrative, which are shared between L1 and L2 writing, supports macro-organization, building paragraph structure, and expressing coherence in L2 narrative texts, but not the use of semantic and lexical means to signal textual relationships. In the argumentative task, no link between ratings of organization and L1 reading was found but participants with higher L1 reading scores used lower number of connectives in their essays. This can be because more proficient L2 writers tend to use fewer connectives and lower number of connectives that are local cohesive devices are also often associated with higher essay ratings (Kojima & Kaneta, 2022).

CONCLUSION

Our study sought to answer the question how L2 and L1 reading, L1 low-level skills, and WM are related to ratings and the linguistic characteristics (productivity, cohesion, lexical sophistication and diversity, syntactic complexity, and accuracy) of narrative and argumentative texts produced by L2 learners. Our findings showed that L2 reading scores were significantly related to raters' perceptions of writing quality as well as to objective measures of linguistic performance. The results also demonstrated that L1 reading is associated with L2 writing through the mediation of L2 reading skills providing support for the Linguistic Threshold Hypothesis (Alderson, 1984). L2 reading ability was found to be significantly related to productivity, grammatical accuracy, lexical diversity and sophistication, and the organization of L2 narrative and argumentative texts highlighting the substantial role of shared L2 grammatical and lexical resources between reading and writing and the importance of reading skills for monitoring and revising L2 written output. Participants with higher L1 reading ability and L1 low-level skills received higher scores for spelling and mechanics as well as organization in the narrative text and used fewer connectives in the argumentative text. These results show that L1 reading experience might influence how coherently and efficiently L2 writers can organize their texts. The narrative texts of students with lower L1 low-level skills contained more spelling errors indicating that students experiencing difficulties with word-level decoding and rapid automated naming in their L1 are likely to face challenges in spelling in writing tasks where they cannot avoid words that they are uncertain how to spell. WM was found to play a limited role in L2 writing performance with significant relationships observed between backward digit span and spelling accuracy in the argumentative task and lexical frequency in the narrative task. These results seem to suggest that for the investigated sample of students the effects of WM capacity might be potentially reduced due to their higher-level proficiency and the observed effects vary as a function of the cognitive demands of the writing tasks (cf. and Kormos', 2023, Task-Mediated Cognitive Model of L2 Writing and Writing to Learn).

From a pedagogical perspective, the findings of our study indicate that L2 learners with lower levels of L1 reading ability and L1 word-level decoding skills and rapid automated naming speed, which might be indicative of literacy-related difficulties such as dyslexia, might experience challenges in different areas of L2 writing. Therefore, in diagnosing potential reasons for L2 writing difficulties, L1 literacy-related factors should also be considered. Our study suggests that L2 writers who struggle with specific aspects of composing such as using sophisticated vocabulary and spelling in certain task types might have weaknesses in WM and L1 low-level skills. These L2 learners could receive targeted support such as additional vocabulary and spelling instruction or focused feedback and more revision opportunities. The results of our research also suggest that developing L1 and L2 reading skills might not only enhance text-comprehension abilities in both languages but can potentially support the development of L2 writing skills.

Our research has several limitations including the size of the sample and the involvement of participants from a well-developed urban area of a European country. Therefore, future research could focus on less industrialized and low-resource and disadvantaged socio-economic contexts in which students might have less well-developed L1 literacy skills and lower L2 proficiency. We only administered two writing tasks and solely focused on the written product and did not investigate L2 writing processes or students' uptake of feedback on their written texts. Our study did not include an assessment of L1 writing skills, which could also have played an important role in our participants' L2 writing performance. In the future, it would also be important to uncover how L1 and L2 reading abilities, L1 writing ability, and L1 low-level skills influence L2 writing processes using concurrent think-aloud, keystroke logging, or eye-tracking technology. The impact of interventions aimed at enhancing L1 and L2 reading skills on L2 writing processes and outcomes would also yield relevant pedagogical insights. Further task types, including integrated reading-to-write and listening-to-write tasks potentially involving multi-modal input and output could also be investigated as their use is becoming increasingly common in instruction and assessment. It would also be useful to explore how L2 learners with different L1 and L2 literacy skills and cognitive profile utilize supportive digital and artificial intelligence tools for their writing and how they process and react to feedback received.

AUTHOR CONTRIBUTIONS

Judit Kormos: Writing—original draft; methodology; conceptualization; data collection; formal analysis. **Csilla Bartha**: Writing—review and editing; conceptualization.

ACKNOWLEDGMENTS

This study was funded under the MTA Distinguished Guest Scientists Fellowship Programme.

DATA AVAILABILITY STATEMENT

Research data are not shared.

ORCID

Judit Kormos https://orcid.org/0000-0002-2643-7222

ENDNOTES

[1] Following Michel et al. (2019), we originally administered a forward digit span test and an emotional Stroop task. However, when we conducted the confirmatory factor analyses, these variables failed to converge on a single factor. A further problem with these measures was that they did not have acceptable psychometric properties (presence of large number of outliers, skewness, and kurtosis values with non-normal distribution).

[2] The argumentative and narrative tasks might also have drawn on different facets of the participants' creativity which was discussed in Kormos et al. (2024) and Kormos and Suzuki (2024).

REFERENCES

Adams, A. M., & Guillot, K. (2008). Working memory and writing in bilingual students. *ITL—International Journal of Applied Linguistics*, *156*, 13–28. https://doi.org/10.2143/ITL.156.0.2034417

Ahmed, Y., Kent, S., Cirino, P. T., & Keller-Margulis, M. (2022). The not-so-simple view of writing in struggling readers/writers. *Reading & Writing Quarterly*, *38*, 272–296. https://doi.org/10.1080/10573569.2021.1948374

Alderson, J. C (1984). Reading in a foreign language: A reading problem or a language problem? In J. C. Alderson & A. H. Urquhart (Eds.), *Reading in a foreign language* (pp. 1–27). Longman.

Alderson, J. C., Haapakangas, E.-L., Nieminen, L., Huhta, A., & Ullakonoja, R. (2015). *The diagnosis of reading in a second or foreign language*. Routledge.

Baddeley, A., & Hitch, G. (1974). Working memory. *The Psychology of Learning and Motivation*, *8*, 47–89. https://doi.org/10.1016/S0079-7421(08)60452-1

Barrouillet, P., Bernardin, S., Portrat, S., Vergauwe, E., & Camos, V. (2007). Time and cognitive load in working memory. *Journal of Experimental Psychology: Learning, Memory, and Cognition*, *33*, 570–585. https://doi.org/10.1037/0278-7393.33.3.570

Bergsleithner, J. M. (2010). Working memory capacity and L2 writing performance. *Ciências & Cogniçao*, *15*, 2–20. https://bit.ly/30JOsrh

Bernhardt, E. B., & Kamil, M. L. (1995). Interpreting relationships between L1 and L2 reading: Consolidating the linguistic threshold and the linguistic interdependence hypotheses. *Applied Linguistics*, *16*, 15–34. https://doi.org/10.1093/applin/16.1.15

Berninger, V. W., & Abbott, R. D. (2010). Listening comprehension, oral expression, reading comprehension, and written expression: Related yet unique language systems in grades 1, 3, 5, and 7. *Journal of Educational Psychology*, *102*, 635–651. https://doi.org/10.1037/a0019319

Berninger, V. W., & Swanson, H. L. (1994). Modifying Hayes and Flower's model of skilled writing to explain beginning and developing writing. In J. S. Carlson, & E. C. Butterfield (Eds.), Advances in cognition and educational practice, Vol. 2. Children's writing: Toward a process theory of the development of skilled writing (pp. 57–81). JAI.

Berninger, V. W., & Winn, W. D. (2006). Implications of advancements in brain research and technology for writing development, writing instruction, and educational evolution. In C. A. MacArthur, S. Graham, & J. Fitzgerald (Eds.), *Handbook of writing research* (pp. 96–114). The Guilford Press.

Blomert, L., & Vaessen, A. (2009). *Differentiaal Diagnostiek van Dyslexie: Cognitieve analyse van lezen en spellen*. Boom Test, Amsterdam, NL.

Booton, S. A., Menon, P., Kyriacou, M., & Murphy, V. A. (2024). Exploring the writing attainment gap: Profiling writing challenges and predictors for children with English as an additional language. *Reading and Writing*, *38*, 843–863. https://doi.org/10.1007/s11145-024-10540-9

Carson, J. E., Carrell, P. L., Silberstein, S., Kroll, B., & Kuehn, P. A. (1990). Reading-writing relationships in first and second language. *TESOL Quarterly*, *24*, 245–266. https://doi.org/10.2307/3586900

Chiappe, P., Siegel, L. S., & Gottardo, A. (2002). Reading-related skills of kindergartners from diverse linguistic backgrounds. *Applied Psycholinguistics*, *23*, 95–116. https://doi.org/10.1017/S014271640200005X

Cho, H., & Brutt-Griffler, J. (2015). Integrated reading and writing: A case of Korean English language learners. *Reading in a Foreign Language*, *27*, 242–261. https://doi.org/10.64152/10125/66886

Coltheart, M. (1981). The MRC psycholinguistic database. *The Quarterly Journal of Experimental Psychology Section A*, *33*, 497–505. https://doi.org/10.1080/14640748108400805

Council of Europe. Council for Cultural Co-operation. Education Committee. Modern Languages Division. (2001). *Common European framework of reference for languages: Learning, teaching, assessment*. Cambridge University Press.

Cummins, J. (1979). Cognitive/academic language proficiency, linguistic interdependence, the optimum age question and some other matters. *Working Papers on Bilingualism*, *19*, 121–129.

Erdos, C., Genesee, F., Savage, R., & Haigh, C. (2014). Predicting risk for oral and written language learning difficulties in students educated in a second language. *Applied Psycholinguistics*, *35*, 371–398. https://doi.org/10.1017/S0142716412000422

Fazio, D., Ferrari, L., Testa, S., Tamburrelli, F., Marra, E., Biancardi, M., & Marzocchi, G. M. (2021). Second-language learning difficulties in Italian children with reading difficulties. *British Journal of Educational Psychology*, *91*, 63–77. https://doi.org/10.1111/bjep.12348

Fitzgerald, J., & Shanahan, T. (2000). Reading and writing relations and their development. *Educational Psychologist*, *35*, 25–37. https://doi.org/10.1207/S15326985EP35015

Geva, E., & Ryan, E. B. (1993). Linguistic and cognitive correlates of academic skills in first and second languages. *Language Learning*, *43*, 5–42. https://doi.org/10.1111/j.1467-1770.1993.tb00171.x

Grabe, W., & Yamashita, J. (2022). L1 and L2 reading relationships. In W. Grabe & J. Yamashita (Eds.), *Reading in a second language: Moving from theory to practice* (pp. 168–194). Cambridge University Press.

Harrison, G. L., Goegan, L. D., Jalbert, R., McManus, K., Sinclair, K., & Spurling, J. (2016). Predictors of spelling and writing skills in first-and second-language learners. *Reading and Writing*, *29*, 69–89. https://doi.org/10.1007/s11145-015-9580-1

Hayes, J. R. (2012). Modelling and remodeling writing. *Written Communication*, *29*, 369–388. https://doi.org/10.1177/0741088312451260

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

Herbert, K. E. D., Massey-Garrison, A., & Geva, E. (2020). A developmental examination of narrative writing in EL and EL1 school children who are typical readers, poor decoders, or poor comprehenders. *Journal of Learning Disabilities*, *53*, 36–47. https://doi.org/10.1177/0022219419881625

Hu, L.-t., & Bentler, P. M. (1998). Fit indices in covariance structure modeling: Sensitivity to underparameterized model misspecification. *Psychological Methods*, *3*, 424–453. https://doi.org/10.1037/1082-989X.3.4.424

Ito, F. (2011). L2 reading–writing correlation in Japanese EFL high school students. *The Language Teacher*, *35*, 23–29. https://doi.org/10.37546/JALTTLT35.5-2

Jeon, E. H., & Yamashita, J. (2022). L2 reading comprehension and its correlates: An updated meta-analysis of correlation coefficients. In E. H. Jeon & Y. In'nami (Eds.), *Understanding L2 proficiency: Theoretical and meta-analytic investigation* (pp. 29–86). John Benjamins. https://doi.org/10.1075/bpa.13.03jeo

Jongejan, W., Verhoeven, L., & Siegel, L. (2007). Predictors of reading and spelling abilities in first-and second-language learners. *Journal of Educational Psychology*, *99*, 835–851. https://doi.org/10.1037/0022-0663.99.4.835

Kellogg, R. T. (1996). A model of working memory in writing. In C. M. Levy & S. Ransdell (Eds.), *The science of writing: Theories, methods, individual differences, and applications* (pp. 57–71). Lawrence Erlbaum.

Kim, K. J., & Pae, T. I. (2021). Examining the simultaneous effects of L1 writing, L2 reading, L2 proficiency, and affective factors on different task types of L2 writing. *Discourse Processes*, *58*, 662–680. https://doi.org/10.1080/0163853X.2021.1872989

Kim, K. J., & Pae, T. I. (2023). The moderating effects of L2 proficiency on the relationship between reading and writing in L2 and across L1 and L2. Foreign Language Annals, *56*, 75–101. https://doi.org/10.1111/flan.12647

Kim, M., Crossley, S. A., & Kim, B. K. (2022). Second language reading and writing in relation to first language, vocabulary knowledge, and learning backgrounds. *International Journal of Bilingual Education and Bilingualism*, *25*, 1992–2005. https://doi.org/10.1080/13670050.2020.1838434

Kim, M., Tian, Y., & Crossley, S. A. (2021). Exploring the relationships among cognitive and linguistic resources, writing processes, and written products in second language writing. *Journal of Second Language Writing*, *53*, 100824. https://doi.org/10.1016/j.jslw.2021.100824

Kim, Y. S. G. (2020). Toward integrative reading science: The direct and indirect effects model of reading. *Journal of Learning Disabilities*, *53*, 469–491. https://doi.org/10.1177/002221942090823

Kim, Y. S. G., & Schatschneider, C. (2017). Expanding the developmental models of writing: A direct and indirect effects model of developmental writing (DIEW). *Journal of Educational Psychology*, *109*, 35–50. https://doi.org/10.1037/edu0000129

Koda, K. (2005). *Insights into second language reading: A cross-linguistic approach*. Cambridge University Press.

Kojima, M., & Kaneta, T. (2022). L2 writing and its internal correlates: A meta-analysis. In E. H. Jeon & Y. In'nami (Eds.), *Understanding L2 proficiency: Theoretical and meta-analytic investigations* (pp. 109–158). John Benjamins.

Kormos, J. (2017). *The second language learning processes of students with specific learning difficulties*. Routledge.

Kormos, J. (2023). The role of cognitive factors in second language writing and writing to learn a second language. *Studies in Second Language Acquisition*, *45*, 622–646. https://doi.org/10.1017/S0272263122000481

Kormos, J. (2025). The role of first language skills, working memory, and anxiety in second language reading: Implications for assessment of language learners with specific learning differences. *Studies in Second Language Acquisition*, 1–23. https://doi.org/10.1017/S0272263125101010

Kormos, J., Kosak-Babuder, M., & Pizorn, K. (2019). The role of low-level first language skills in second language reading, reading-while-listening and listening performance: A study of young dyslexic and non-dyslexic language learners. *Applied Linguistics*, *40*, 834–858. https://doi.org/10.1093/applin/amy028

Kormos, J., & Sáfár, A. (2008). Phonological short-term memory, working memory and foreign language performance in intensive language learning. *Bilingualism*, *11*, 261–271. https://doi.org/10.1017/S1366728908003416

Kormos, J., & Suzuki, S. (2024). Creativity and the linguistic features of argumentative and narrative written task performance. *System*, *127*, 103531. https://doi.org/10.1016/j.lindif.2024.102500

Kormos, J., Suzuki, S., & Rossi, O. (2024). The role of creativity in second language writing performance. *Learning and Individual Differences*, *114*, 102500. https://doi.org/10.1016/j.lindif.2024.102500

Landerl, K., Ramus, F., Moll, K., Lyons, I. M., Ansari, D., & White, S. (2013). Predictors of developmental dyslexia in European orthographies with varying complexity. *Journal of Child Psychology and Psychiatry*, *54*, 686–694. https://doi.org/10.1111/jcpp.12029

Lee, J. (2020). Effects of linguistic and affective variables on middle school students' writing performance in the context of English as a foreign language: An approach using structural equation modelling. *Reading and Writing*, *33*, 1235–1262. https://doi.org/10.1007/s11145-019-10007-2

Lee, S. (2005). Facilitating and inhibiting factors in English as a foreign language writing performance: A model testing with structural equation modeling. *Language Learning*, *55*, 335–374. https://doi.org/10.1111/j.0023-8333.2005.00306.x

Li, S. (2023). Working memory and second language writing: A systematic review. *Studies in Second Language Acquisition*, *45*, 647–679. https://doi.org/10.1017/S0272263123000189

Li, C., Wei, L., & Lu, X. (2024). Task complexity and L2 writing performance of young learners: Contributions of cognitive and affective factors. *Modern Language Journal*, *108*, 741–770. https://doi.org/10.1111/modl.12954

Łockiewicz, M., & Jaskulska, M. (2016). Difficulties of Polish students with dyslexia in reading and spelling in English as L2. *Learning and Individual Differences*, *51*, 256–264. https://doi.org/10.1016/j.lindif.2016.08.037

Manchón, R. M., McBride, S., Mellado Martínez, M. D., & Vasylets, O. (2023). Working memory, L2 proficiency, and task complexity: Independent and interactive effects on L2 written performance. *Studies in Second Language Acquisition*, *45*, 737–764. https://doi.org/10.1017/S0272263123000141

Mavrou, I. (2020). Working memory, executive functions, and emotional intelligence in second language writing. *Journal of Second Language Writing*, *50*, 100758. https://doi.org/10.1016/j.jslw.2020.100758

McCutchen, D. (1996). A capacity theory of writing: Working memory in composition. *Educational Psychology Review*, *8*, 299–324. https://doi.org/10.1007/BF01464076

Michel, M., Kormos, J., Brunfaut, T., & Ratajczak, M. (2019). The role of working memory in young second language learners' written performances. *Journal of Second Language Writing*, *45*, 31–45. https://doi.org/10.1016/j.jslw.2019.03.002

Mizumoto, A., Shintani, N., Sasaki, M., & Teng, M. F. (2024). Testing the viability of ChatGPT as a companion in L2 writing accuracy assessment. *Research Methods in Applied Linguistics*, *3*, https://doi.org/10.1016/j.rmal.2024.100116

Mujtaba, S. M., Kamyabi Gol, A., & Parkash, R. (2021). A study on the relationship between language aptitude, vocabulary size, working memory, and L2 writing accuracy. *Foreign Language Annals*, *54*, 1059–1081. https://doi-org.ezproxy.lancs.ac.uk/10.1111/flan.12584

Pae, T. I. (2019). A simultaneous analysis of relations between L1 and L2 skills in reading and writing. *Reading Research Quarterly*, *54*, 109–124. https://doi.org/10.1002/rrq.216

Pae, T.-I. (2018). Effects of task type and L2 proficiency on the relationship between L1 and L2 in reading and writing. *Studies in Second Language Acquisition*, *40*, 63–90. https://doi.org/10.1017/S0272263116000462

Palladino, P., Cismondo, D., Ferrari, M., Ballagamba, I., & Cornoldi, C. (2016). L2 spelling errors in Italian children with dyslexia. *Dyslexia*, *22*, 158–172. https://doi.org/10.1002/dys.1522

Peng, A., Orosco, M. J., Wang, H., Swanson, H. L., & Reed, D. K. (2021). Cognition and writing development in early adolescent English learners. *Journal of Educational Psychology*, *114*, 1136–1155. https://doi.org/10.1037/edu0000695

Plonsky, L., & Oswald, F. L. (2014). How big is "big"? Interpreting effect sizes in L2 research. *Language Learning*, *64*, 878–912. https://doi.org/10.1111/lang.12079

Révész, A., Michel, M., & Lee, M. (2017). Investigating IELTS academic Writing Task 2: Relationships between cognitive writing processes, text quality, and working memory. *IELTS Research Reports*, *3*, 5–44. https://ielts.org/researchers/our-research/research-reports/investigating-ielts-academic-writing-task-2-relationships-between-cognitive-writing-processes-text-quality-and-working-memory

R Core Team, R. (2020). R Core Team R: a language and environment for statistical computing. *Foundation for Statistical Computing*.

Rosseel, Y. (2012). lavaan: An R Package for Structural Equation Modeling. *Journal of Statistical Software*, *48*, 1–36. https://doi.org/10.18637/jss.v048.i02

Sasaki, M., & Hirose, K. (1996). Explanatory variables for EFL students' expository writing. *Language Learning*, *46*, 93–127. https://doi.org/10.1111/j.1467-1770.1996.tb00641.x

Schoonen, R. (2022). L2 writing: Theory and research. In E. H. Jeon & Y. In'nami (Eds.), *Understanding L2 proficiency: Theoretical and meta-analytic investigations* (pp. 87–107). John Benjamins Publishing Company.

Sehlström, P., Waldmann, C., Steinvall, A., & Levlin, M. (2022). Swedish (L1) and English (L2) argumentative writing of upper secondary students with reading difficulties. *L1-Educational Studies in Language and Literature*, *22*, 1–22. https://doi.org/10.21248/L1ESLL.2022.22.1.405

Sparks, R. L., & Ganschow, L. (1993). The impact of native language learning problems on foreign language learning: Case study illustrations of the linguistic coding deficit hypothesis. *Modern Language Journal*, *77*, 58–74. https://doi.org/10.2307/329559

van Gelderen, A., Schoonen, R., de Glopper, K., Hulstijn, J., Simis, A., Snellings, P., & Stevenson, M. (2004). Linguistic knowledge, processing speed, and metacognitive knowledge in first- and second language reading comprehension: A componential analysis. *Journal of Educational Psychology*, *96*, 19–30. https://doi.org/10.1037/0022-0663.96.1.19

van Setten, E. R. H., Tops, W., Hakvoort, B. E., van der Leij, A., Maurits, N. M., & Maassen, B. A. M. (2017). L1 and L2 reading skills in Dutch adolescents with a familial risk of dyslexia. *PeerJ*, *5*, e3895. https://doi.org/10.7717/peerj.3895

Vasylets, O., & Marín, J. (2021). The effects of working memory and L2 proficiency on L2 writing. *Journal of Second Language Writing*, *52*, 100786. https://doi.org/10.1016/j.jslw.2020.100786

Wade-Woolley, L., & Siegel, L. S. (1997). The spelling performance of ESL and native speakers of English as a function of reading skill. *Reading and Writing*, *9*, 387–406. https://doi.org/10.1023/A:1007902323536

Yoon, H. J., & Polio, C. (2017). The linguistic development of students of English as a second language in two written genres. *TESOL Quarterly*, *51*, 275–301. https://doi.org/10.1002/tesq.296

Zabihi, R. (2018). The role of cognitive and affective factors in measures of L2 writing. *Written Communication*, *35*, 32–57. https://doi.org/10.1177/0741088317735836

Zhang, H. (2021). The longitudinal effect of morphological awareness on higher-order literacy skills among college L2 learners. *Contemporary Educational Psychology*, *65*, Article 101969. https://doi.org/10.1016/j.cedpsych.2021.101969

SUPPORTING INFORMATION

Additional supporting information can be found online in the Supporting Information section at the end of this article.

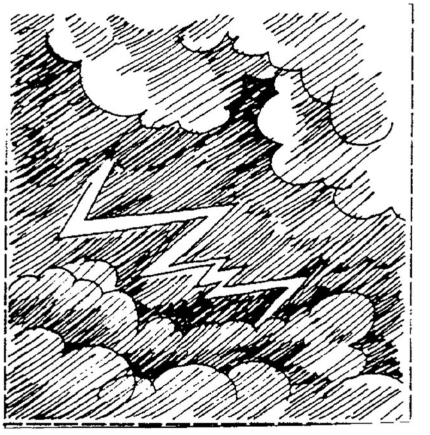
How to cite this article: Kormos, J., & Bartha, C. (2026). The role of first and second language reading, first language low-level skills, and working memory in second language writing. *Modern Language Journal*, 1–29. https://doi.org/10.1111/modl.70030

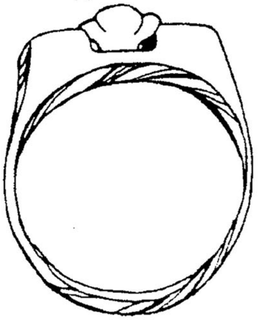
APPENDIX A: ARGUMENTATIVE TASK

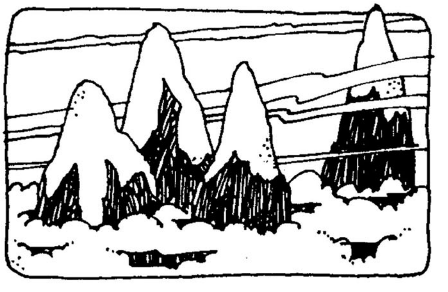
You saw a post online that said that ***Teenagers are too young to teach other people about anything***. Write a response in which you express your disagreement with this statement. Express your opinion and support your arguments with examples. Write about 200–250 words. You can use this paper to plan your essay, but please type the essay on the online interface. You have 20–25 minutes to write the response.

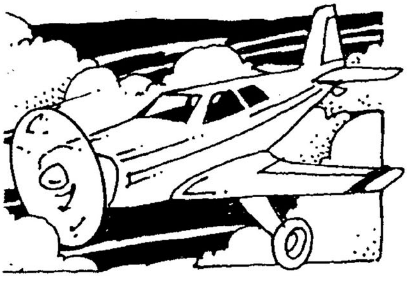
NARRATIVE TASK

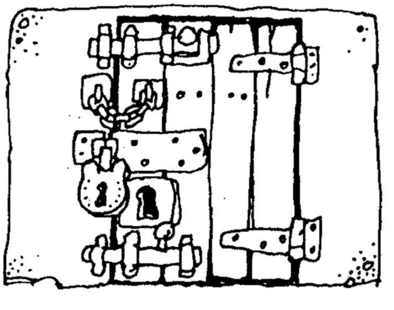
You can see six pictures below. Your task is to write a story which includes all the elements shown in the pictures. You must use all the pictures in the story, but you may also add extra information if you wish. Write about 150–200 words. You can use this paper to plan your story, but please type the story on the online interface. You have 20–25 minutes to write the response.











APPENDIX B: THE RATING SCALE FOR WRITING TASKS

	Content					
	Argumentative	Narrative	Organization	Vocabulary	Language use	Spelling and mechanics
4	• The student's position is **clear** • Arguments are **well developed** with **sufficient** support	• A fully **developed** storyline • Main actors, sequence of events and setting **described in great detail**	• **Well** organized • **Logical** sequencing • Cohesive	• **Sophisticated** range • **Effective** word/ idiom choice and usage, word form mastery • **Appropriate** register	• **Effective complex** constructions • **Few** errors	**Mastery** of conventions **Few** errors of spelling, punctuation, capitalization, paragraphing
3	• The student's position is **generally evident** • Arguments are **appropriately developed** with **some level** of support	• An **appropriately developed** storyline • Main actors, sequence of events, and setting **described but not in much detail**	**Loosely** organized but main ideas stand out **Logical but incomplete** sequencing	**Adequate** range **Occasional** errors of word/idiom form, choice, usage Meaning **not obscured**	**Effective but simple** constructions **Minor problems** in complex constructions **Several** errors Meaning **seldom obscured**	**Occasional** errors of spelling, punctuation, capitalization, paragraphing Meaning **not obscured**
2	• The student's position is **somewhat evident** • Arguments are **not developed in detail** and support for them is **not adequate**	• The events are **described** • **Lack of detail** about actors, sequence, and setting	Ideas **confused** or **disconnected** **Lacks logical** sequencing and development	• **Limited** range • **Frequent** errors of word/idiom form, choice, usage • Meaning **confused or obscured**	**Major problems** in simple/complex constructions **Frequent** errors and/or fragments, run-ons, deletions Meaning **confused or obscured**	**Frequent** errors of spelling, punctuation, capitalization, paragraphing Meaning **confused or obscured**
1	• The student's position is **rarely evident** • Arguments are **weak** and support for claims is **not provided**	• The story is **underdeveloped** • Information about key events, sequence, and context is **missing**	**No** organization OR **not enough to evaluate**	**Little knowledge** of English vocabulary, idioms, word form OR **not enough to evaluate**	Virtually **no mastery** of sentence construction rules **Dominated** by errors, does not communicate OR **not enough to evaluate**	• **No mastery** of conventions • **Dominated** by errors of spelling, punctuation, capitalization, paragraphing • OR **not enough to evaluate**

APPENDIX C

TABLE C1 Inter-rater reliability of the writing scores.

Variable	Spearman's ρ	p
Argumentative content	0.754	<.001
Argumentative organization	0.699	<.001
Argumentative lexis	0.794	<.001
Argumentative language use	0.747	<.001
Argumentative mechanics	0.700	<.001
Story content	0.855	<.001
Story organization	0.807	<.001
Story lexis	0.842	<.001
Story languague use	0.828	<.001
Story mechanics	0.843	<.001

TABLE C2 Confirmatory factor analysis results for the latent variables of L2 writing, L2 reading, L1 reading and L1 low-level skills.

Task	df	χ^2	p-value	χ^2/df ratio	CFI	TLI	SRMR	RMSEA [90% CI]
L2 writing—argumentative	5	25.047	<.001	5.00	0.891	0.901	0.061	0.078 [0.056, 0.105]
L2 writing—narrative	5	20.753	.001	5.18	0.918	0.837	0.049	0.074 [0.043, 0.111]
L2 reading	3	47.661	<.001	15.87	1.000	1.00	0.000	0.000 [0.000, 0.108]
L1 low-level skills	6	97.264	<.001	16.21	1.00	1.03	0.015	0.000 [0.000, 0.126]

Abbreviations: CFI, comparative fit index; RMSEA, root mean square error of association; SRMR, standardized root mean square; TLI, Tucker–Lewis index.
The cutoff values for good fit: χ^2/df ratio < 2.0; SRMR < .08; CFI and TLI > .90, RMSEA < .08.

TABLE C3 Descriptive statistics for task performance measures in the argumentative task.

	Min	Max	Mean	SD
Productivity factor	−2.41	3.54	.00	1.00
Semantic overlap factor	−3.06	3.04	.00	1.00
Frequency of all connectives	0.01	0.12	.08	.020
Frequency of logical connectives	0.01	0.11	.06	.018
Frequency of causal connectives	0.01	0.09	.02	.010
Lexical diversity factor	−4.53	2.41	.00	1.00
Lexical frequency factor	−2.91	3.03	.00	1.00
Lexical retrieval speed factor	−2.47	2.97	.00	1.00
MRC familiarity—content words	576	597	586.49	4.37
Hypernymy verbs	1	2	1.70	.154
Mean length of T-unit	9	64	17.67	6.989
Syntactic dependency factor	−2.79	2.84	.00	1.00
Frequency of grammatical errors/100 words	.00	.19	.05	.04
Frequency of spelling errors/100 words	.00	.11	.02	.02

Note: Factor scores are regression scores that have a mean of 0 and an SD of 1.

TABLE C4 Descriptive statistics for task performance measures in the narrative task.

	Min	Max	Mean	SD
Productivity factor	−2.17	4.07	.00	1.00
Semantic overlap factor	−5.06	4.40	.00	1.00
Frequency of all connectives	.01	.10	.06	.01
Frequency of logical connectives	.01	.08	.03	.01
Frequency of causal connectives	.00	.05	.01	.01
Lexical diversity factor	−3.58	3.06	.00	1.00
Lexical frequency factor	−2.93	2.43	.00	1.00
Lexical retrieval speed factor	−2.27	2.23	.00	1.00
MRC familiarity -content words	566.19	596.33	583.51	6.03
Hypernymy verbs	1.19	1.94	1.57	.13
Mean length of T-unit	7.77	90.00	14.26	10.33
Syntactic dependency factor	−8.92	1.50	.00	1.00
Frequency of grammatical errors/100 words	.00	.23	.05	.03
Frequency of spelling errors/100 words	.00	.18	.02	.02

Note: Factor scores are regression scores that have a mean of 0 and an SD of 1.00.

TABLE C5 Factor loadings of linguistic variables.

Factor	Variables	Argumentative	Narrative
Productivity	Number of content words	.962	.972
	Number of function words	.931	.967
	Number of clauses	.779	.849
	Number of T-units	.857	.825
Cohesion: Semantic overlap	Synonym overlap sentences (nouns)	.735	.776
	Latent semantic analysis (sentence)	.833	.898
	Word2vec similarity (sentence)	.800	.459
Lexical diversity	D-index all word types	.865	.864
	D-index function words	.872	.849
	MTLD	.861	.892
Lexical sophistication: Frequency	COCA Academic Frequency Log Function Words	.952	.933
	COCA Academic Frequency Log Content Words	.741	.749
	COCA Academic Bigram Frequency Log	.792	.713
Lexical retrieval speed	Lexical decision time Z score content words	.818	.882
	Lexical decision time SD content words	.656	.734
	Word naming response time Z score content words	.797	.895
	Word naming response time SD score content words	.705	.689
Syntactic complexity:	Dependents per clause	.775	.920
	Adverbial modifiers per clause	.816	.676
	Average construction frequency—academic (log transformed)	.766	.681