International Journal of TESOL Studies (2025)
251107, 1-18 https://doi.org/10.58304/ijts.251107

Article

The Mediating Role of Vocabulary in Linking L1 Reading and L2 Proficiency

Siwon Park*
Tetsuya Fukuda
Paul Wadden
Juntendo University, Japan

Received: 20 August, 2025/Received in revised form: 29 October, 2025/Accepted: 10 November, 2025/Available online: 28 November, 2025

Abstract
This study examined how first language (L1) reading relates to second language (L2) proficiency and whether L2 vocabulary mediates that relationship in a Japanese university context. Participants were 213 first-year students at a private university. L1 reading was assessed with the Reading Skills Test (RST), L2 vocabulary with the New General Service List Test (NGSLT), and L2 proficiency with the TOEFL ITP sections of Listening, Structure and Written Expression, and Reading. A structural equation model estimated both direct and indirect paths among the three constructs. Results showed a positive effect of L1 reading on L2 vocabulary ($\beta = .28$, $p < .05$, 95% CI [.12, .44]) and a strong effect of vocabulary on L2 proficiency ($\beta = .58$, $p < .001$, 95% CI [.45, .71]). The direct path from L1 reading to L2 proficiency was small and not significant ($\beta = .06$, $p > .05$), indicating full mediation through vocabulary. Measurement limitations, including weak factor loadings for some TOEFL subtests ($\lambda = .49$-$.64$), warrant cautious interpretation. The findings highlight vocabulary as the primary channel through which L1 literacy advantages influence the placement-oriented proficiency measures used here. Pedagogically, sustained vocabulary instruction and continued support for L1 reading appear essential. Future research should broaden the definition of proficiency to include productive skills and adopt a longitudinal design to test whether vocabulary growth precedes later gains in overall proficiency.

Keywords
L1 reading ability, L2 proficiency, L2 vocabulary, structural equation modeling (SEM)

1 Introduction

Reading is a complex cognitive process that involves decoding written symbols, retrieving lexical knowledge, and constructing meaning from text (Castles et al., 2018; Duke & Cartwright, 2021). For experienced native language readers, this process is typically straightforward and automatic. However,

*Corresponding author. Email: s.park.ll@juntendo.ac.jp

second language (L2) learners must acquire new linguistic structures and learn unfamiliar vocabulary and possibly different writing systems. This raises important questions about which first language (L1) reading skills transfer to L2 learning contexts and how this transfer affects overall communication, particularly regarding vocabulary.

The relationship between L1 reading ability and L2 acquisition is multifaceted (Peng et al., 2020). L1 reading skills may provide a solid foundation for developing effective reading skills in L2, potentially facilitating faster language acquisition. The effectiveness of this transfer depends on the linguistic distance between languages and may operate indirectly through L2 vocabulary knowledge.

Prior studies have explored various aspects of these relationships. Some of them have shown that L1 and L2 reading comprehension processes are positively correlated (e.g., Bernhardt & Kamil, 1995; Lesaux & Siegel, 2003), indicating that reading skills can transfer across languages. There is considerable evidence that L1 reading skills can be applied in the L2 (Durgunoğlu, 2002; Goodrich & Lonigan, 2017). Another line of research emphasizes the importance of vocabulary knowledge for L2 proficiency (e.g., Nation, 2013) and illustrates that lexis is a crucial factor in both receptive and productive language abilities.

However, few studies have investigated the interplay between L1 reading ability, L2 proficiency, and L2 vocabulary at the tertiary level in Japan within a single, comprehensive model. Previous research has examined the interplay of L1 reading, L2 proficiency, and vocabulary across diverse learner populations (e.g., Bernhardt & Kamil, 1995; Durgunoğlu, 2002; Jeon & Yamashita, 2022; Lesaux & Siegel, 2003; Sparks et al., 2012). These studies provide evidence of cross-linguistic transfer and underscore the importance of vocabulary for L2 outcomes. However, few have investigated these relationships in the context of Japanese learners at the tertiary level using a unified SEM framework. This study addresses both this contextual and methodological gap.

In this study, 'L2 proficiency' refers to listening, grammar, and reading comprehension as measured by the TOEFL ITP. 'L1 reading ability' is defined as the capacity to accurately comprehend written Japanese texts, assessed with the Reading Skills Test (RST). 'L2 vocabulary knowledge' refers to the breadth of lexical knowledge as measured by the New General Service List Test (NGSLT).

2 Literature Review

2.1 First language reading proficiency and its foundational elements

First-language (L1) reading proficiency is a multi-component construct that integrates phonological awareness, orthographic and morphological knowledge, syntactic processing, and metacognitive control (Grabe & Jiang, 2018; Zhao et al., 2016). Skilled reading emerges from the interaction of these resources, enabling efficient decoding, inferencing, and comprehension monitoring. Accordingly, when we refer to L1 reading "supporting" L2 learning below, we mean that the same component skills (lexical knowledge, lexical fluency, and metacognitive control) operate in both languages. We do not claim that L1 reading performance transfers wholesale; rather, the skills that enable it do. While certain components are language-specific, such as phonological decoding shaped by the writing system, others are more transferable across languages. Higher-order strategies such as inferencing and comprehension monitoring, as well as metalinguistic awareness, have been shown to influence second-language (L2) outcomes (Durgunoğlu, 2002; Goodrich & Lonigan, 2017).

L1–L2 connections are best explained by shared cognitive and linguistic foundations rather than a one-way transfer of reading ability. Skills that underlie efficient L1 reading, including vocabulary breadth, lexical access speed (fluency), and metacognitive monitoring, also underpin reading development in a second language (Durgunoğlu, 2002; Goodrich & Lonigan, 2017). In this view, L1 reading predicts L2 outcomes because both rely on the same component skills.

In L2 contexts, vocabulary depth and lexical fluency are strong predictors of reading performance, highlighting that the quality and accessibility of word knowledge matter as much as sheer size (Tong et al., 2023). This supports a shared-mechanisms account based on common lexical and strategic resources rather than a simple directional effect from L1 to L2.

This overlap also resonates with applied work in the International Journal of TESOL Studies, which advocates task-supported reading instruction integrating vocabulary and strategy training (Ellis, 2024; Jung, 2024), a practical route to strengthening the foundational elements relevant to both L1 and L2 reading.

2.2 Influence of L1 reading on L2 learning

The shared cognitive-linguistic foundations outlined above provide a basis for understanding how L1 literacy relates to L2 learning. Although L1 reading skills offer this foundation, their influence is shaped by proficiency thresholds and language distance.

According to Cummins' linguistic threshold hypothesis (1976, 1979), a certain level of L2 knowledge is required before L1 reading skills can be applied productively. In contrast, the interdependence hypothesis (Cummins, 1979) assumes that both languages draw on a common underlying proficiency encompassing cognitive and linguistic skills developed through literacy experiences. Empirical findings support both perspectives: Bernhardt and Kamil (1995) and Lesaux and Siegel (2003) reported robust cross-linguistic correlations, while Pae (2018) showed that transfer effects varied by task type and proficiency level.

Recent longitudinal research further supports this interdependence framework while emphasizing shared component skills rather than direct transfer. Raudszus et al. (2021) found that vocabulary breadth, grammatical knowledge, and situation-model construction jointly predicted comprehension outcomes in both L1 and L2 reading. Their results point to reciprocal influences among underlying literacy skills, consistent with the view that common cognitive and linguistic resources, rather than discrete language-specific abilities, explain cross-linguistic relationships.

Orthographic distance also constrains transfer. Akamatsu (2003) showed that Japanese learners of English face particular decoding challenges compared with alphabetic L1 readers, limiting straightforward transfer. More recent neurocognitive and behavioral evidence supports this claim: Dong et al. (2021) demonstrated that orthographic transparency affects cross-language pattern adaptation, and Kuperman (2025) reported that L1–L2 language distance reliably predicted English reading fluency and comprehension after controlling for component skills. Together, these findings suggest that while L1 reading skills can support L2 learning, their influence is conditional on learners' L2 proficiency and the linguistic distance between languages. These converging results underscore vocabulary knowledge as a central interface between L1 and L2 literacy, setting the stage for the next section.

2.3 Vocabulary as a mediating mechanism

Vocabulary is widely regarded as a central determinant of L2 proficiency and a likely mediator of the L1–L2 relationship (Nation, 2013). Without sufficient lexical coverage, learners cannot effectively employ higher-order comprehension strategies, regardless of L1 reading ability. Empirical evidence consistently supports this view. Sparks et al. (2012) found that L1 reading achievement and print exposure predicted later L2 proficiency, with vocabulary growth serving as a likely pathway. Similarly, Jeon and Yamashita's (2022) meta-analysis showed that vocabulary knowledge accounted for a substantial proportion of variance in L2 reading comprehension.

Recent syntheses further confirm vocabulary's central role. Zhang and Zhang's (2022) meta-analysis estimated the correlation between L2 vocabulary and reading at .57 and found a comparable link with

listening. Yan et al. (2021) reported a strong vocabulary-reading association among Chinese learners, moderated by educational stage. At the construct level, Tong et al. (2023) demonstrated that receptive depth was more strongly related to reading than breadth, and that vocabulary fluency moderated both relationships. Longitudinal evidence also supports the mediating role of vocabulary. Shibasaki et al. (2015) found that L1 literacy influenced L2 comprehension indirectly through vocabulary and working memory among Japanese high school students.

While most studies, including the present one, assess breadth of vocabulary knowledge, research increasingly emphasizes the importance of depth (semantic networks, collocations, and multiword expressions) and fluency (speed of access). Webb et al.'s (2023) meta-analysis of incidental learning quantified how vocabulary develops through input, clarifying how reading exposure contributes to the mediator identified in the present model. Recent longitudinal analyses further support vocabulary's mediating role (Teng, 2025; Teng & Cui, 2025a, 2025b). Together, these findings underscore vocabulary's role as the conduit through which L1 literacy advantages manifest in L2 proficiency.

2.4 Evidence from Japanese learners

Although L1–L2 transfer has been widely studied, relatively few investigations have modeled these processes among Japanese learners, particularly at the tertiary level. Yamashita (2002) found that L1 reading ability and L2 proficiency could compensate for one another in predicting L2 comprehension, while Yamashita (2011) demonstrated links between reading attitudes across languages. Shibasaki et al. (2015) used SEM with Japanese high school students to show that vocabulary mediated the relationship between L1 literacy and L2 comprehension. More recently, Jeon and Yamashita (2022) reported a moderate correlation between L1 and L2 reading among Japanese undergraduates.

Other work in Japan has clarified the structure of vocabulary knowledge and its role in comprehension. Koizumi and In'nami (2020) used both conventional and Bayesian SEM with adult learners to show that vocabulary size and depth are distinct yet highly correlated constructs. Ueno and Takeuchi (2022) modeled self-regulated vocabulary learning among Japanese high school students, confirming vocabulary's central role in comprehension. Longitudinal studies of reading rate development among Japanese adolescents further demonstrated that vocabulary growth sets the stage for later comprehension gains (Yamashita et al., 2023).

Collectively, these studies reinforce the view that vocabulary functions as a mediating mechanism linking L1 literacy and L2 comprehension. They also establish the Japanese context as a valuable setting for examining how shared literacy skills manifest in L2 learning outcomes.

2.5 Methodological precedents for SEM

Structural equation modeling (SEM) offers a powerful approach for examining complex relationships among L1 reading, L2 vocabulary, and L2 proficiency. Unlike simple correlations, SEM enables researchers to model latent constructs, estimate direct and indirect effects, and account for measurement error (Kline, 2023). Numerous studies in applied linguistics have demonstrated its value. Pae (2018) used SEM to examine how task type and proficiency moderated the L1–L2 relationship, while Shibasaki et al. (2015) modeled indirect effects of L1 literacy through vocabulary and working memory. Koizumi and In'nami (2020) further showed SEM's capacity to distinguish related but distinct dimensions of vocabulary knowledge.

At a broader level, meta-analytic SEM studies of L2 reading have consistently shown that comprehension-related skills, which are strongly associated with vocabulary, account for more variance in reading outcomes than decoding skills (Hamada, 2024; Peng et al., 2020). These findings highlight the value of SEM for testing whether L1 reading effects on L2 proficiency are direct or operate primarily through vocabulary and justify its use in the present study.

2.6 Summary and gap

The reviewed literature indicates that L1 reading provides a cognitive and linguistic foundation for L2 learning, but its influence is conditioned by language distance, proficiency thresholds, and learners' lexical resources. Vocabulary knowledge consistently emerges as the key mechanism linking L1 literacy to broader L2 proficiency, and recent studies emphasize the need to distinguish among vocabulary breadth, depth, and fluency. Within the Japanese context, evidence supports vocabulary's mediating role, yet relatively few investigations have examined these relationships comprehensively at the tertiary level using SEM.

To address this gap, the present study examines the interrelationships among L1 reading ability, L2 vocabulary knowledge, and L2 proficiency. Specifically, it aims to clarify whether and how vocabulary mediates the connection between L1 reading and L2 proficiency.

Methodologically, we employ SEM to test a hypothesized model linking three latent constructs: L1 reading, L2 vocabulary, and L2 proficiency (Figure 1). L1 reading is measured by six subskills of the Reading Skills Test, L2 vocabulary by five levels of the New General Service List Test (NGSLT), and L2 proficiency by three sections of the TOEFL ITP (Listening, Structure and Written Expression, and Reading). The structural portion specifies a mediation pathway in which L1 reading predicts L2 vocabulary, and L2 vocabulary predicts L2 proficiency, with a direct path from L1 reading to L2 proficiency estimated to test whether the relationship involves full or partial mediation. The model is overidentified and estimated with a marker-variable approach, with residuals constrained to be uncorrelated unless theoretically justified.

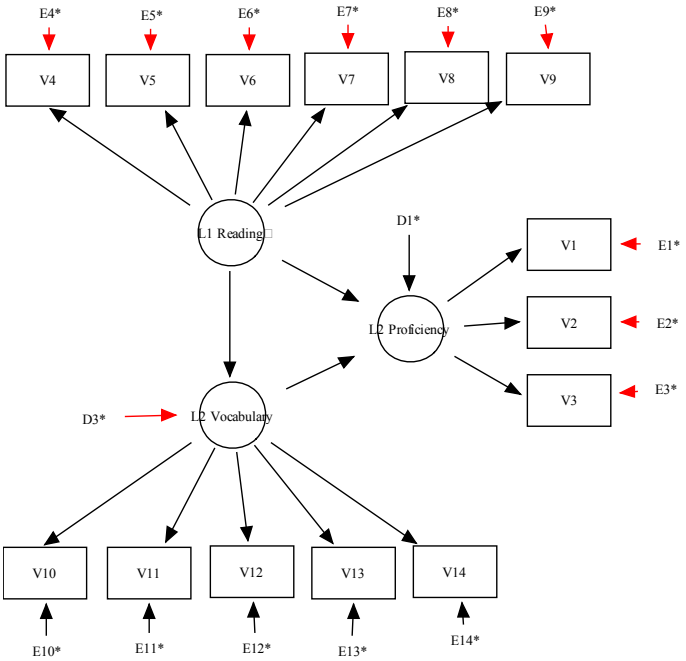
SEM is particularly well-suited to this task because it models latent variables, tests direct and indirect effects simultaneously, and accounts for measurement error, providing a comprehensive view of how shared literacy skills may influence L2 proficiency through vocabulary.

The present study is guided by the following research questions (RQs):

- RQ1: To what extent does L1 reading ability predict L2 proficiency?
- RQ2: To what extent does L2 vocabulary mediate the relationship between L1 reading ability and L2 proficiency?

Figure 1

The Hypothesized Model



Note. V1: Listening Comprehension; V2: Structure and Written Expression; V3: Reading Comprehension; V4: Dependency Parsing; V5: Coreference Resolution; V6: Paraphrase Detection; V7: Inference; V8: Image Identification; V9: Example Identification; V10: Level 1; V11: Level 2; V12: Level 3; V13: Level 4; V14: Level 5.

3 Method

3.1 Participants

The analytic sample consisted of 213 first-year undergraduates at a private university in eastern Japan. Of these, 124 were female (58%) and 89 were male (42%). The participants' mean age was 18.6 (*SD* = 0.7; range 18–20). All were enrolled in a liberal arts program that required English coursework and a second foreign language of their choice.

The larger incoming cohort included 257 students; 44 were excluded because they self-reported a non-Japanese first language on the TOEFL answer sheet. This exclusion was consistent with the study's focus on Japanese L1 reading ability.

Testing and data collection were conducted during the first semester of the 2024 academic year, within the students' orientation period. All participants were informed that their anonymized test scores might be used for research purposes, and they were given the opportunity to opt out without penalty.

3.2 Instruments

3.2.1 L2 proficiency

L2 English proficiency was assessed using the TOEFL ITP, which all first-year students complete before matriculation for placement into proficiency-banded English classes. Scores from the three sections were used: Listening Comprehension, Structure and Written Expression, and Reading Comprehension. Testing took place in large lecture halls under standardized conditions, approximately two weeks before the semester began.

Item-level data were unavailable from the local ETS office, precluding sample-specific reliability estimation. Publisher documentation reports internal consistency coefficients of .88–.93 for the three sections (ETS, 2023). While these estimates indicate strong internal reliability, they may not fully represent the current sample.

The measure primarily reflects receptive skills and grammatical knowledge, and does not include speaking or writing components. Consequently, the proficiency construct used in this study should be interpreted as a receptive-grammar composite rather than a comprehensive measure of communicative competence.

3.2.2 L1 reading ability

L1 reading ability was assessed with the Reading Skills Test (RST), a diagnostic measure developed by Arai and colleagues to evaluate expository text comprehension as a set of reading processes rather than topic knowledge (Arai et al., 2017). The instrument includes six domains that target different aspects of comprehension: Dependency Parsing, Coreference Resolution, Paraphrase Detection, Inference, Image Identification, and Example Identification (Arai et al., 2017). Items are delivered online using an item response theory-based computerized adaptive format that tailors difficulty to the examinee and aims to maintain precision across the ability range (Arai et al., 2017; Reckase, 2000; Weiss, 1982).

Participants completed the adaptive RST individually in a computer lab. Scores were reported on a standardized IRT scale (M = 50, SD = 10). Because item-level data were not retained, internal consistency for this sample could not be estimated; previous studies report acceptable reliability (Arai et al., 2017).

3.2.3 L2 vocabulary

L2 vocabulary knowledge was measured using the New General Service List Test (NGSLT; Stoeckel & Bennett, 2015), which is based on the New General Service List (NGSL) (Browne, 2014; Browne et al., 2013). The NGSL identifies high-frequency English words essential for comprehension and instruction, updating the original General Service List (West, 1953).

The English-only version of NGSLT Form A was adopted in this study. It contains five levels with 20 multiple-choice items per level (total = 100 items) that assess receptive knowledge of high-frequency vocabulary. The test was administered online via Google Forms without time limits, and scoring was automated upon submission.

Internal-consistency reliability coefficients (Cronbach's α) for this sample were .67, .76, .74, .75, and .79 for Levels 1–5, respectively. Consistent with prior findings, Level 1 showed marginal reliability due to guessing effects, while the upper levels demonstrated satisfactory precision. The NGSLT assesses vocabulary breadth rather than depth or productive knowledge, which should be considered when interpreting the results.

3.3 Procedures

Data were collected during the orientation period preceding the start of the academic year. All tests were administered under standardized proctoring conditions by trained staff. The sequence was as follows:

- TOEFL ITP (administered approximately two weeks before the semester began) in paper format for placement into English proficiency bands.
- NGSLT (administered during the orientation week) conducted online during English orientation sessions and automatically scored upon submission.
- RST (administered during the freshman orientation period) completed individually under supervision in a computer lab.

Participation involved routine institutional assessments, and no additional testing was required for research purposes. After data collection, all personally identifying information was removed, and each participant was assigned a numeric code. Analyses were conducted using anonymized datasets stored on an encrypted university server.

3.4 Analysis

Data were analyzed using IBM SPSS 29.0 (IBM Corp., 2022) and EQS 6.4 for Windows (Bentler, 2018). Descriptive statistics and zero-order correlations were computed first. Using the covariance structure derived from the observed scores, we estimated model-fit indices and factor loadings for the specified models. The full variance–covariance matrix used for estimation is provided in Appendix A.

Structural equation modeling (SEM) was then used to estimate direct and indirect relationships among latent variables, control for measurement error, and test the hypothesized mediation of vocabulary (Kline, 2023). The model was estimated with the robust maximum-likelihood (MLR) method, which adjusts standard errors and fit statistics for moderate non-normality. Missing data (< 2%) were handled

using full-information maximum likelihood (FIML), which utilizes all available data points without listwise deletion.

To evaluate the mediating (indirect) effects, we employed a non-parametric bootstrapping procedure implemented in EQS 6.4. Bias-corrected bootstrap confidence intervals were generated based on 5,000 resamples, and standard errors were obtained from the empirical bootstrap distribution. An indirect path was interpreted as statistically significant if its 95% confidence interval did not include zero. This resampling approach provides a robust test of mediation by avoiding the normal-theory assumption of indirect-effect sampling distributions, offering greater accuracy than the Sobel test in moderately sized samples.

Model fit was assessed using χ^2/df, CFI, TLI, and RMSEA, with cutoff criteria of CFI and TLI $\geq .90$ and RMSEA $\leq .08$. With a sample size of 213, the analysis met conventional SEM guidelines of 5–10 cases per estimated parameter (Boomsma, 1982; Kline, 2023).

3.5 Ethical considerations

Participation involved standard institutional testing (TOEFL ITP, NGSLT, RST). Students were informed that anonymized results might be used for research purposes, and they could decline participation without disadvantage. All data were de-identified prior to analysis and stored securely on an encrypted server.

4 Results

4.1 Descriptive and reliability statistics

Table 1 reports descriptive statistics for all measures for the five NGSLT levels. The means and standard deviations indicate adequate dispersion without floor or ceiling effects. Because the local ETS office conducted TOEFL ITP scoring, item-level data were not available, and internal consistency for this sample could not be estimated. Publisher documentation reports high internal consistency for the three TOEFL ITP sections ($\alpha = .88–.93$; ETS, 2023)

Item-level data for the RST were also unavailable because it is a commercial computerized adaptive test. Although adaptive delivery typically improves measurement precision across the ability range (Reckase, 2000; Weiss, 1982), the absence of raw responses prevents sample-specific reliability analysis.

For the NGSLT, internal consistency by level was $\alpha = .67, .76, .74, .75$, and .79 for Levels 1 through 5, respectively, with Level 1 best characterized as marginal. This pattern mirrors earlier validation studies reporting lower reliability for the easiest items, which are more susceptible to guessing (Stoeckel & Bennett, 2015). These scores fall within the expected range for first-year Japanese university students, whose receptive vocabulary typically averages around 3,500–4,000 word families (McLean et al., 2014). Analyses that involve the RST and NGSLT Level 1 are interpreted with this constraint in mind.

Table 1

Descriptive Statistics for Study Measures

Measure	*M*	*SD*	Min.	Max.
L2 Proficiency (TOEFL ITP)				
LC	43.98	3.71	31	61
GR	38.97	4.81	31	53
RC	42.09	5.18	31	54

L1 Reading (RST)				
L1R1	51.70	7.32	31	70
L1R2	51.70	9.20	0	75
L1R3	50.67	8.79	0	77
L1R4	50.99	7.49	32	68
L1R5	52.26	7.98	0	73
L1R6	52.45	7.58	30	69
L2 Vocabulary (NGSLT)				
Level 1	16.98	2.29	8	20
Level 2	16.07	3.14	2	20
Level 3	15.67	3.04	0	20
Level 4	13.10	3.23	0	20
Level 5	12.72	3.80	0	20

Note. $N = 213$. LC = Listening Comprehension; GR = Structure and Written Expression; RC = Reading Comprehension.

4.2 Convergent and discriminant validity

In SEM, inferences about relations among latent variables are only as credible as the measurement model that defines those latent variables. Convergent validity asks whether a set of indicators shares sufficient common variance to warrant a single construct; discriminant validity asks whether the constructs are empirically distinct. Weak convergence typically attenuates structural paths and indirect effects, while poor discriminant validity can introduce multicollinearity that inflates or even reverses coefficients and obscures mediation. For these reasons, both forms of validity are examined before interpreting the structural paths (Kline, 2023).

Table 2

Standardized Factor Loadings (λ) for Individual Measures

Factors	**Measures**	**Standardized loading (λ)**
L2 Proficiency (TOEFL ITP)	Listening Comprehension	.49
	Structure and Written Expression	.49
	Reading Comprehension	.64
L1 Reading Skills (RST)	Dependency Parsing	.57
	Coreference Resolution	.62
	Paraphrase Detection	.73
	Inference	.64
	Image Identification	.74
	Example Identification	.74
L2 Vocabulary (NGSLT)	Level 1	.56
	Level 2	.77
	Level 3	.81
	Level 4	.80
	Level 5	.83

Convergent validity was assessed using standardized factor loadings and average variance extracted (AVE), with .60 for loadings and .50 for AVE as practical benchmarks. As shown in Figure 2 and Table 2, eight of 14 indicators met the .60 loading threshold, with RST indicators ranging from .57 to .74 and NGSLT Levels 2–5 ranging from .77 to .83. Four indicators fell below .60: TOEFL ITP Listening Comprehension (.49), Structure and Written Expression (.49), RST Dependency Parsing (.57), and NGSLT Level 1 (.56). These patterns suggest weaker convergent evidence for the L2 Proficiency latent and only marginal support for the two noted subtests; results are interpreted with appropriate caution.

Discriminant validity was evaluated with the Fornell–Larcker criterion by comparing the square root of each construct's AVE with its latent correlations (Table 3). The square roots of AVE were .55 for L2 Proficiency, .68 for L1 Reading, and .76 for L2 Vocabulary. Each exceeded the corresponding inter-construct correlations (largest correlations: L2 Proficiency–L2 vocabulary = .42, L1 Reading–L2 Vocabulary = .26, L2 Proficiency–L1 Reading = .15), indicating that the three constructs are empirically separable within the measurement model. For completeness, the AVE values were .30, .46, and .58, respectively, which points to limited convergent validity for L2 Proficiency and marginal convergent validity for L1 Reading. Taken together, these findings indicate that while the overall measurement model is adequate, the latent L2 Proficiency factor should be interpreted conservatively given its weaker loadings on Listening and Structure subtests.

Table 3

Discriminant Validity Assessment Using the Fornell–Larcker Criterion

	L2P (√AVE)	***L1R (√AVE)***	***L2V (√AVE)***
L2 Proficiency (L2P) (√AVE)	**.55**	.15	.42
L1 Reading (L1R) (√AVE)	.15	**.68**	.26
L2 Vocabulary (L2V) (√AVE)	.42	.26	**.76**

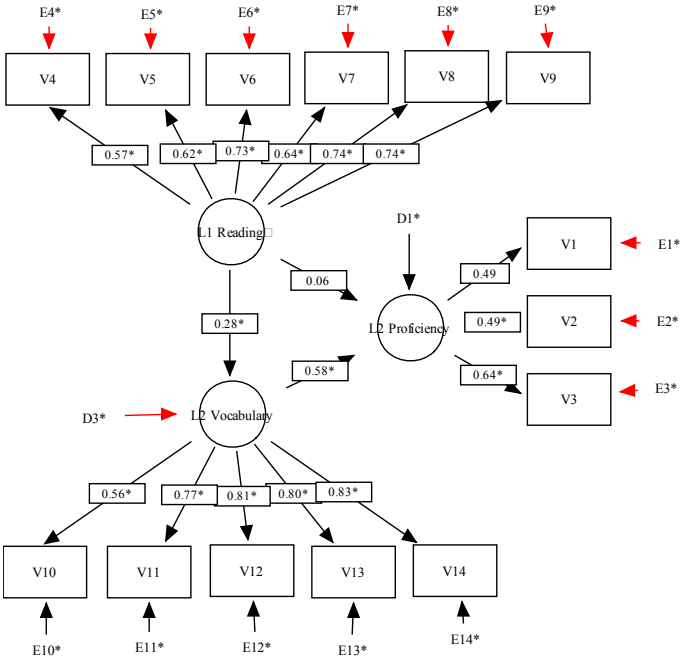
Note. Diagonal entries are the square roots of average variance extracted (√AVE). Off-diagonals are latent correlations (φ). Each √AVE exceeds the absolute correlations in its row and column, indicating discriminant validity. $N = 213$.

4.3 Model identification and fit

We specified a structural model with three latent constructs: L2 proficiency, L1 reading, and L2 vocabulary. These were measured by 14 observed indicators. Each latent factor was scaled using a marker-variable approach (one loading fixed to 1.0). Residuals were constrained to be uncorrelated unless theory justified otherwise. The model is overidentified ($df = 74$) and therefore testable. Model estimation was based on the observed variance–covariance matrix (Appendix A) using robust maximum-likelihood estimation (MLR). Global fit was acceptable, $\chi^2(74) = 136.122$, $\chi^2/df = 1.84$, CFI = .94, TLI = .93, RMSEA = .06. These indices meet common benchmarks for acceptable model-data correspondence (CFI and TLI ≥ .90, RMSEA ≤ .08; Kline, 2023). Given the sample size of 213, these fit statistics indicate that the hypothesized three-factor structure provides a reasonable approximation of the observed data.

Figure 2

Hypothesized Measurement Model with Standardized Factor Loadings



Note. Latent variables are shown as circles and observed indicators as rectangles. One loading per latent was fixed to 1.00 for scale identification. Standardized loadings are displayed on each path. V1 = Listening Comprehension; V2 = Structure and Written Expression; V3 = Reading Comprehension; V4 = Dependency Parsing; V5 = Coreference Resolution; V6 = Paraphrase Detection; V7 = Inference; V8 = Image Identification; V9 = Example Identification; V10 = Level 1; V11 = Level 2; V12 = Level 3; V13 = Level 4; V14 = Level 5. E1–E14 and D1–D3 denote residuals. $N = 213$.

4.4 Structural model results

The structural model (Figure 2) revealed three key pathways. L1 reading positively predicted L2 vocabulary (β = .28, SE = .12, p = .02, 95% CI [.05, .51]), representing a small-to-medium effect (Cohen, 1988). L2 vocabulary, in turn, strongly predicted L2 proficiency (β = .58, SE = .09, $p < .001$, 95% CI [.41, .75]), indicating a large effect. The direct path from L1 reading to L2 proficiency was small and non-significant (β = .06, SE = .12, p = .62, 95% CI [-.17, .29]).

The indirect effect of L1 reading on L2 proficiency through vocabulary was significant (β = .16, SE = .07, p = .02, 95% CI [.03, .30]), indicating that vocabulary fully mediated the relationship between L1 reading and L2 proficiency. L2 vocabulary accounted for 34% of the variance in L2 proficiency (R^2 = .34).

These results align with the shared-skills framework proposed in Section 2.1, suggesting that the cognitive and lexical processes supporting L1 reading contribute to L2 development primarily through vocabulary growth rather than direct skill transfer. The absence of a significant direct path may also reflect language distance effects between Japanese and English (Akamatsu, 2003; Kuperman, 2025) and the receptive emphasis of the TOEFL ITP measure, which limits its sensitivity to higher-level comprehension processes. Together, these findings provide empirical support for a fully mediated model in which L2 vocabulary serves as the central conduit linking L1 reading and L2 proficiency.

5 Discussion

We tested a hypothesized structural model linking L1 reading, L2 vocabulary, and L2 proficiency in a first-year university cohort. The model yielded a small, non-significant direct path from L1 reading to L2 proficiency (β = .06, $p > .05$), a positive path from L1 reading to L2 vocabulary (β = .28, $p < .05$), and a strong path from L2 vocabulary to L2 proficiency (β = .58, $p < .001$). The sections below interpret these findings with respect to the research questions and situate them within prior work.

5.1 RQ1: Does L1 reading ability predict L2 proficiency?

The direct effect of L1 reading on L2 proficiency was negligible and non-significant (β = .06, 95% CI [-.17, .29]), indicating that any L1 reading advantage operates through alternative pathways rather than broadly enhancing proficiency. The result refines earlier reports of positive links between L1 and L2 reading and of cross-linguistic transfer of core reading resources (Bernhardt & Kamil, 1995; Durgunoğlu, 2002; Goodrich & Lonigan, 2017; Lesaux & Siegel, 2003). Studies with Japanese undergraduates have also documented correlational ties between L1 and L2 reading (Jeon & Yamashita, 2022). Our estimates indicate that, at university entry, such influence does not appear as a general proficiency boost once vocabulary knowledge is taken into account.

Two strands from the literature help explain this pattern. Threshold accounts propose that L1 reading resources become usable in L2 only after learners possess sufficient L2 knowledge to support those processes (Cummins, 1976, 1979; Yamashita, 2002, 2011). If many students are early in their university studies, conditions for a sizable direct L1–L2 proficiency link may not yet be present. Orthographic distance can further constrain transfer. Readers moving from a non-alphabetic L1 to English face additional decoding and mapping challenges, which makes an indirect route through language-specific knowledge more plausible at early stages (Akamatsu, 2003; Kuperman et al., 2025).

Considered together, these perspectives align with a shared-skills interpretation: cognitive and lexical processes developed through L1 literacy matter, but their influence is contingent on intermediate L2 resources, particularly vocabulary. Measurement decisions in the present study also limit the scope of inference, since proficiency was defined by receptive skills on the TOEFL ITP, a test in which vocabulary coverage strongly influences performance.

5.2 RQ2: Does L2 vocabulary mediate the relation between L1 reading and L2 proficiency?

The model supported a mediated relation. Stronger L1 reading predicted larger L2 vocabularies (β = .28, $p < .05$), and vocabulary, in turn, was closely tied to proficiency (β = .58, $p < .001$). This large effect indicates that a one standard deviation increase in L2 vocabulary corresponds to more than half a standard deviation increase in L2 proficiency, highlighting vocabulary's central role in language development. This finding supports vocabulary as a limiting factor for L2 comprehension and performance (Lervåg & Aukrust, 2010; Nation, 2013).

Prior longitudinal and cross-sectional work links L1 literacy experiences and print exposure to later L2 vocabulary and reading, which then relate to broader outcomes (Jeon & Yamashita, 2022; Sparks et al., 2012). Studies that model vocabulary as a distinct construct similarly identify a bridge from L1 literacy to L2 comprehension through lexis (Shibasaki et al., 2015). In a first-year Japanese context, our estimates fit that trajectory: L1 reading appears to set conditions that make L2 word learning more efficient, and once vocabulary size supports adequate text coverage, gains are detectable on the listening and reading tasks that shape placement-oriented proficiency.

The mediated account is also consistent with interdependence and threshold explanations. Interdependence claims that skills developed in L1 can support L2 learning given adequate exposure and motivation (Kim et al., 2020), while threshold formulations specify that vocabulary is one of the

resources that must reach workable levels before transfer is productive in L2 reading contexts (Cummins, 1976, 1979; Lee & Schallert, 2014). Evidence that lexical skills themselves can transfer across languages further explains why vocabulary growth functions as a natural conduit in this linkage (Toplu & Erten, 2023; Tong et al., 2023).

Mechanistically, extensive L1 reading is associated with faster form-meaning mapping, flexible inferencing, and effective metacognitive control during reading; these processes should accelerate L2 word learning, which then supports the proficiency outcomes sampled here (Nation, 2013; Lervåg & Aukrust, 2010).

5.3 Implications for theory and practice

Taken together, the findings point to an integrated view in which interdependence and threshold ideas both hold but become visible at different stages. Early in university, L1 reading does not directly raise a general proficiency index once vocabulary is modeled ($\beta = .06$, $p > .05$); influence is exerted through lexical growth, which then supports broader L2 performance ($\beta = .58$, $p < .001$).

The pattern helps reconcile mixed reports in the literature. Studies that detect direct L1 effects often involve more advanced learners or longer exposure, where vocabulary thresholds have likely been crossed and broader transfer can surface (Bernhardt & Kamil, 1995; Lesaux & Siegel, 2003). The present results, therefore, align with a shared-skills framework in which the same cognitive-linguistic abilities that support L1 literacy also facilitate L2 development, primarily via vocabulary.

For instruction, the model underscores the importance of sustaining L1 reading habits while making L2 vocabulary growth a central, explicit target through rich input, deliberate study of high-utility words and multiword expressions, spaced retrieval, and extensive reading at appropriate levels, practices long emphasized in applied linguistics (Nation, 2013; Ellis, 2024; Jung, 2024). Integrating such approaches may accelerate learners' progress toward the proficiency levels required for effective academic and professional communication. In institutional contexts, vocabulary diagnostics could be triangulated with proficiency testing through validated academic-vocabulary measures such as those developed by Nguyen and Hoang (2024), supporting data-informed curriculum and placement decisions.

6 Conclusion

6.1 Summary of findings

The present study examined how L1 reading, L2 vocabulary, and L2 proficiency are related among first-year university students in Japan. The hypothesized structural model showed acceptable fit and a consistent pattern: the direct path from L1 reading to L2 proficiency was small and not significant ($\beta = .06$, $p > .05$), L1 reading predicted L2 vocabulary ($\beta = .28$, $p < .05$), and L2 vocabulary strongly predicted L2 proficiency ($\beta = .58$, $p < .001$). Together, these results indicate full mediation, with vocabulary functioning as the principal conduit between L1 literacy and L2 performance.

Interpreted within a shared-skills framework, the findings suggest that cognitive and lexical processes developed through L1 reading contribute to L2 learning primarily by facilitating vocabulary growth rather than by directly enhancing general proficiency. This account complements interdependence and threshold perspectives and underscores vocabulary's pivotal role in cross-linguistic literacy development.

6.2 Limitations and directions for future research

Several constraints should frame the interpretation of these findings. Measurement limitations posed the greatest challenges. Four indicators showed factor loadings below .60, including both TOEFL

Listening and Structure subtests (λ = .49 each), and the L2 Proficiency latent variable demonstrated weak convergent validity (AVE = .30). Internal consistency could not be estimated for TOEFL ITP and RST because item-level data were unavailable, though publisher documentation reports adequate reliability. These limitations likely attenuated path estimates and may have contributed to the fully mediated pattern observed here.

Construct coverage was also limited. Proficiency was defined by receptive subtests of the TOEFL ITP (Listening, Structure and Written Expression, and Reading), leaving productive skills unassessed. Consequently, the findings apply mainly to receptive proficiency rather than comprehensive communicative ability.

The cross-sectional design and single-institution sample further limit causal inference and generalizability. Longitudinal and multi-group designs could clarify developmental trajectories, threshold effects, and potential variability across proficiency levels or institutional contexts.

Future research should extend this work in four directions. First, broaden the proficiency construct to include productive skills and integrated tasks, testing whether direct L1 reading effects emerge when proficiency is more fully represented. Second, employ longitudinal and multi-group SEM to capture developmental changes and subgroup differences. Third, strengthen measurement by adding reliable indicators, reporting ω alongside α, and considering composite specifications when latent loadings are weak. Future studies might also employ validated academic-vocabulary instruments that capture lexical depth and domain-specific use, such as the test validated through Rasch analysis by Nguyen and Hoang (2024). Fourth, distinguish vocabulary breadth, depth, and fluency, include multiword expressions, and test robustness with bootstrapped indirect effects and measurement invariance analyses.

Advancing along these lines would yield a more comprehensive account of how L1 literacy experiences shape L2 learning and would help identify pedagogical practices that help learners draw on L1 reading strengths to support L2 development.

Appendix A

Covariance Matrix Used in Model Estimation

	V1	V2	V3	V4	V5	V6	V7	V8	V9	V10
V1	13.7966									
V2	5.2776	23.1030								
V3	6.1579	6.8800	26.8025							
V4	-2.1082	2.4013	-0.7851	53.6170						
V5	2.8152	4.8844	3.6211	26.5965	84.5755					
V6	-2.6413	6.6561	4.1803	27.9879	32.8608	77.2258				
V7	-2.7648	5.4555	4.4792	19.0339	25.3867	34.8238	56.0849			
V8	0.8265	7.6287	6.2126	23.6107	30.0531	41.8416	25.9636	63.6844		
V9	0.3346	6.2034	7.0700	22.7892	38.1676	32.0177	26.9125	33.7311	57.4941	
V10	1.6174	3.2399	2.2805	3.3511	4.6647	4.7636	5.3967	4.7529	4.7564	5.2494
V11	1.8460	4.0822	5.0264	3.2543	6.3448	4.7626	6.1595	6.0075	5.9647	3.8743
V12	2.3177	2.6133	5.2197	0.8078	1.3640	3.6281	4.2534	4.6426	3.4121	2.7139
V13	2.5449	2.3756	4.5563	0.9393	2.8812	0.1141	3.1823	1.8088	2.4740	3.2477
V14	3.1160	4.7420	7.7153	1.9669	5.6289	2.9331	6.1363	6.3238	5.1047	3.7811

	V11	V12	V13	V14
V11	9.8488			
V12	6.1223	9.2594		
V13	5.8701	6.5577	10.4421	
V14	7.1850	7.7419	8.4773	14.4486

References

Akamatsu, N. (2003). The effects of first language orthographic features on second language reading in text. *Language Learning, 53*(2), 207–231. https://doi.org/10.1111/1467-9922.00216

Arai, N. H., Todo, N., Arai, T., Bunji, K., Sugawara, S., Inuzuka, M., Matsuzaki, T., & Ozaki, K. (2017). Reading Skill Test to diagnose basic language skills in comparison to machines. *CogSci 2017: Proceedings of the 39th Annual Meeting of the Cognitive Science Society*, 1556–1561.

Bentler, P. M., & Wu, E. J. C. (2018). *EQS 6.4 for Windows* [Computer software]. Multivariate Software, Inc.

Bernhardt, E. B., & Kamil, M. L. (1995). Interpreting relationships between L1 and L2 reading: Consolidating the linguistic threshold and interdependence hypotheses. *Applied Linguistics, 16*(1), 15–34. https://doi.org/10.1093/applin/16.1.15

Boomsma, A. (1982). The robustness of LISREL against small sample sizes in factor analysis models. In K. Jöreskog & H. Wold (Eds.), *Systems under indirect observation* (Part I, pp. 149–173). North-Holland.

Browne, C. (2014). A new general service list: The better mousetrap we've been looking for? *Vocabulary Learning and Instruction, 3*(2), 1–10. https://doi.org/10.7820/vli.v03.2.browne

Browne, C., Culligan, B., & Phillips, J. (2013). *A new general service list*. https://www.newgeneralservicelist.com/new-general-service-list

Castles, A., Rastle, K., & Nation, K. (2018). Ending the reading wars: Reading acquisition from novice to expert. *Psychological Science in the Public Interest, 19*(1), 5–51. https://doi.org/10.1177/1529100618772271

Cohen, J. (1988). *Statistical power analysis for the behavioral sciences* (2nd ed.). Lawrence Erlbaum Associates.

Cummins, J. (1976). The influence of bilingualism on cognitive growth: A synthesis of research findings and explanatory hypotheses. *Working Papers on Bilingualism, 9*, 1–43.

Cummins, J. (1979). Linguistic interdependence and the educational development of bilingual children. *Review of Educational Research, 49*(2), 222–251. https://doi.org/10.3102/00346543049002222

Dong, J., Li, A., Chen, C., Qu, J., Jiang, N., Sun, Y., Hu, L., & Mei, L. (2021). Language distance in orthographic transparency affects cross-language pattern similarity between native and non-native languages. *Human Brain Mapping, 42*(4), 893–907. https://doi.org/10.1002/hbm.25266

Duke, N. K., & Cartwright, K. B. (2021). The science of reading progresses: Communicating advances beyond the simple view of reading. *Reading Research Quarterly, 56*(S1), 25–44. https://doi.org/10.1002/rrq.411

Durgunoğlu, A. Y. (2002). Cross-linguistic transfer in literacy development and implications for language learners. *Annals of Dyslexia, 52*(1), 189–204. https://doi.org/10.1007/s11881-002-0012-y

Ellis, R. (2024). Task-based and Task-supported Language Teaching. *International Journal of TESOL Studies, 6*(4), 1-13. https://doi.org/10.58304/ijts.20240401

ETS. (2023). *TOEFL ITP: Test and data summary*. ETS.

Goodrich, J. M., & Lonigan, C. J. (2017). Language-independent and language-specific aspects of early literacy: An evaluation of the Common Underlying Proficiency model. *Journal of Educational Psychology, 109*(6), 782–793. https://doi.org/10.1037/edu0000179

Grabe, W., & Jiang, X. (2018). First language and second language reading. In *The TESOL Encyclopedia of English Language Teaching* (pp. 1–7). John Wiley & Sons. https://doi.org/10.1002/9781118784235.eelt0476

Hamada, A., Shimizu, H., Hoshino, Y., Takaki, S., & Ushiro, Y. (2024). Robust evidence for the simple view of second language reading: Secondary meta-analysis of Jeon and Yamashita (2022). *Studies in Second Language Acquisition, 46*(5), 1355–1372. https://doi.org/10.1017/S0272263124000226

IBM Corp. (2022). *IBM SPSS Statistics for Windows* (Version 29.0) [Computer software]. IBM Corp.

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

Jung, J. (2024). Reading Tasks as a Pathway for Embracing TBLT in Asia: A Response to Ellis (2024). *International Journal of TESOL Studies, 6*(4), 55–61. https://doi.org/10.58304/ijts.20240406

Kim, M., Crossley, S. A., & Kim, B. K. (2020). Second language reading and writing in relation to first language, vocabulary knowledge, and learning backgrounds. *International Journal of Bilingual Education and Bilingualism, 25*(6), 1992–2005. https://doi.org/10.1080/13670050.2020.1838434

Kline, R. B. (2023). *Principles and practice of structural equation modeling* (5th ed.). Guilford Press.

Koizumi, R., & In'nami, Y. (2020). Structural equation modeling of vocabulary size and depth using conventional and Bayesian methods. *Frontiers in Psychology, 11*, 618. https://doi.org/10.3389/fpsyg.2020.00618

Kuperman, V. (2025). How does language distance affect reading fluency and comprehension in English as second language? *Studies in Second Language Acquisition*, 1–17. https://doi.org/10.1017/S027226312510065X

Kuperman, V., Schroeder, S., Acartürk, C., Agrawal, N., Alexandre, D. M., Bolliger, L. S., … Siegelman, N. (2025). New data on text reading in English as a second language: The Wave 2 expansion of the Multilingual Eye-Movement Corpus (MECO). *Studies in Second Language Acquisition*, *47*(2), 677–695. https://doi.org/10.1017/S0272263125000105

Lee, J. W., & Schallert, D. L. (1997). The relative contribution of L2 language proficiency and L1 reading ability to L2 reading performance: A test of the threshold hypothesis in an EFL context. *TESOL Quarterly, 31*(4), 713–739. https://doi.org/10.2307/3587757

Lee, J., & Schallert, D. L. (2014). Literate actions, reading attitudes, and reading achievement: Interconnections across languages for adolescent learners of English in Korea. *The Modern Language Journal, 98*(2), 553–573. https://doi.org/10.1111/modl.12088

Lervåg, A., & Aukrust, V. G. (2010). Vocabulary knowledge is a critical determinant of the differences in reading comprehension growth between first and second language learners. *Journal of Child Psychology and Psychiatry, 51*(5), 612–620. https://doi.org/10.1111/j.1469-7610.2009.02185.x

Lesaux, N. K., & Siegel, L. S. (2003). The development of reading in children who speak English as a second language. *Developmental Psychology, 39*(6), 1005–1019. https://doi.org/10.1037/0012-1649.39.6.1005

McLean, S., Hogg, N., & Kramer, B. (2014). Estimations of Japanese university learners' English vocabulary sizes using the Vocabulary Size Test. *Vocabulary Learning and Instruction, 3*(2), 47–55. https://doi.org/10.7820/vli.v03.2.mclean.et.al

Nation, I. S. P. (2013). *Learning vocabulary in another language* (2nd ed.). Cambridge University Press.

Nguyen, T. T. M., & Hoang, T. T. H. (2024). Validating Pecorari, Shaw and Malmström's (2019) Academic Vocabulary Test – Form 1: Evidence from Rasch-based analyses and retrospective focus-group interviews. *International Journal of TESOL Studies, 6*(2), 24–42. https://doi.org/10.58304/ijts.20240203

Pae, T.-I. (2018). Effects of task type and L2 proficiency on the relationship between L1 and L2 in reading and writing: An SEM approach. *Studies in Second Language Acquisition, 40*(1), 63–90. https://doi.org/10.1017/S0272263116000462

Peng, P., Lee, K., Luo, J., Li, S., Joshi, R. M., & Tao, S. (2020). Simple view of reading in Chinese: A one-stage meta-analytic structural equation modeling. *Review of Educational Research, 91*(1), 3–33. https://doi.org/10.3102/0034654320964198

Raudszus, H., Segers, P. C. J., & Verhoeven, L. (2021). Patterns and predictors of reading comprehension growth in first and second language readers. *Journal of Research in Reading, 44*(2), 400–417. https://doi.org/10.1111/1467-9817.12347

Reckase, M. D. (2000). The evolution of computerized testing: Comparing CATs to paper-and-pencil tests. In W. J. van der Linden & C. A. W. Glas (Eds.), *Computerized adaptive testing: Theory and practice* (pp. 31–52). Kluwer Academic.

Shibasaki, H., Tokimoto, S., Ono, Y., Inoue, T., & Tamaoka, K. (2015). English reading comprehension by Japanese high school students: Structural equation modeling including working memory and L1 literacy. *Open Journal of Modern Linguistics, 5*(5), 443–458. https://doi.org/10.4236/ojml.2015.55039

Sparks, R. L., Patton, J., Ganschow, L., & Humbach, N. (2009). Long-term cross-linguistic transfer of skills from L1 to L2. *Language Learning, 59*(1), 203–243. https://doi.org/10.1111/j.1467-9922.2009.00504.x

Sparks, R. L., Patton, J., Ganschow, L., & Humbach, N. (2012). Do L1 reading achievement and L1 print exposure contribute to the prediction of L2 proficiency? *Language Learning, 62*(2), 473–505. https://doi.org/10.1111/j.1467-9922.2012.00694.x

Stoeckel, T., & Bennett, P. (2015). A test of the New General Service List. *Vocabulary Learning and Instruction, 4*(1), 1–8.

Teng, M. F. (2025). Longitudinal analysis of breadth of vocabulary knowledge in mediating metacognitive knowledge and reading in foreign language young learners. *Reading in a Foreign Language, 37*(2), 18-38. https://doi.org/10.64152/10125/67496

Teng, M. F., & Cui, Y. (2025a). The role of vocabulary knowledge, morphological awareness, and working memory in reading comprehension. *Reading in a Foreign Language, 37*(1), 1-21. https://doi.org/10.64152/10125/67481

Teng, M. F., & Cui, Y. (2025b). The contribution of vocabulary knowledge and morphological awareness to reading comprehension in a foreign language. *Journal of Research in Reading, 48*(2), 131-152. https://doi.org/10.1111/1467-9817.70000

Tong, Y., Hasim, Z., & Abdul Halim, H. (2023). The relationship between L2 vocabulary knowledge and reading proficiency: The moderating effects of vocabulary fluency. *Humanities and Social Sciences Communications, 10*, 555. https://doi.org/10.1057/s41599-023-02058-2

Toplu, N. C., & Erten, I. H. (2023). A study into the interplay between first and second language reading motivation, reading habits, and vocabulary size. *Sustainable Multilingualism, 22*(1), 94–137. https://doi.org/10.2478/sm-2023-0005

Ueno, S., & Takeuchi, O. (2022). Self-regulated vocabulary learning in a Japanese high school EFL environment: A structural equation modeling approach. *JACET Journal, 66*, 97-111. https://doi.org/10.32234/jacetjournal.66.0_97

Webb, S., Uchihara, T., & Yanagisawa, A. (2023). How effective is second language incidental vocabulary learning? A meta-analysis. *Language Teaching, 56*(2), 161–180. https://doi.org/10.1017/S0261444822000507

Weiss, D. J. (1982). Improving measurement quality and efficiency with adaptive testing. *Applied Psychological Measurement, 6*(4), 473–492. https://doi.org/10.1177/014662168200600408

West, M. (1953). *A general service list of English words*. Longman.

Yamashita, J. (2002). Mutual compensation between L1 reading ability and L2 language proficiency in L2 reading comprehension. *Journal of Research in Reading, 25*(1), 81–95. https://doi.org/10.1111/1467-9817.00160

Yamashita, J. (2011). The relationship of reading attitudes between L1 and L2: An investigation of adult EFL learners in Japan. *TESOL Quarterly, 41*(1), 81–105. https://doi.org/10.1002/j.1545-7249.2007.tb00041.x

Yamashita, J., Kusanagi, K., & Shiotsu, T. (2023). Development of EFL reading rate in adolescents in Japan. *Journal of Second Language Studies, 6*(2), 238-265. https://doi.org/10.1075/jsls.22005.yam

Yan, M., Li, Y., Sun, X., Zhou, X., Hui, Y., & Li, H. (2021). The roles of decoding and vocabulary in Chinese reading development: Evidence from a 3-year longitudinal study. *The British journal of educational psychology*, *91*(1), 300–314. https://doi.org/10.1111/bjep.12365

Zhang, S., & Zhang, X. (2022). The relationship between vocabulary knowledge and L2 reading/listening comprehension: A meta-analysis. *Language Teaching Research, 26*(4), 696-725. https://doi.org/10.1177/1362168820913998

Zhao, J., Quiroz, B., Dixon, L. Q., & Joshi, R. M. (2016). Comparing Bilingual to Monolingual Learners on English Spelling: A Meta-analytic Review. *Dyslexia (Chichester, England), 22*(3), 193–213. https://doi.org/10.1002/dys.1530

Siwon Park is a professor in the Faculty of International Liberal Arts at Juntendo University, Japan. His research interests include educational measurement, learner characteristics, second language acquisition, global education, interdisciplinary research, and research methodology.

ORCID: https://orcid.org/0009-0004-5561-4162

Tetsuya Fukuda is a senior associate professor in the Faculty of International Liberal Arts at Juntendo University. His research interests include second language motivation, vocabulary test analysis, language program development, and World Englishes.

Paul Wadden is a specially appointed professor in the Faculty of International Liberal Arts at Juntendo University. His research interests include academic English, vocabulary, standardized English proficiency tests (e.g., TOEFL), and curriculum design for English-medium instruction.