

## NGSL Vocabulary Development in Relation to TOEIC Reading Gains

Jean-Pierre J. Richard  
*The University of Nagano*

### Reference Data:

Richard, J.-P. J. (2026). NGSL vocabulary development in relation to TOEIC reading gains. In B. Lacy, M. Swanson, & H. Yoder (Eds.), *LanguageS: Learning, Teaching, Assessing – JALT 50 Years – Challenges and Perspectives*. JALT. <https://doi.org/10.37546/JALTPCP2025-09>

The purpose of this study was to examine the extent to which New General Service List (NGSL) development is associated with Test of English for International Communication (TOEIC) Reading gains within a coordinated two-year English Language Program. Data from one cohort of management majors ( $N = 174$ ) were used to investigate relationships between Year-1 NGSL level progression and TOEIC Reading gains over one (T2–T1) and two (T4–T1) years. Pearson correlations and multiple regression analyses were conducted using baseline proficiency measures and vocabulary development variables. Results showed that NGSL level gain was positively associated with TOEIC Reading gains across both timeframes, even after controlling for baseline proficiency. Initial vocabulary level was also positively associated with later TOEIC Reading gains, whereas higher initial TOEIC Reading scores were associated with smaller gains. These findings suggest that structured, frequency-based vocabulary development may represent one meaningful component of improvement in TOEIC Reading performance in university EFL contexts.

本研究では、2年間の英語教育プログラム(ELP)において、New General Service List(NGSL)の習熟度向上がTOEICリーディング得点の伸びとどの程度関連しているかを検討した。経営学専攻の学部生1学年( $N = 174$ )のデータを用い、1年次におけるNGSLレベルの進展と、1年間(T2–T1)および2年間(T4–T1)のTOEICリーディング得点の伸びとの関係を分析した。分析には、初期英語力指標および語彙発達変数を用いた相関分析と重回帰分析を実施した。その結果、NGSLレベルの伸びは、初期能力を統制した後も、両期間におけるTOEICリーディング得点の伸びと正の関連を示した。また、初期語彙レベルもその後のTOEICリーディング得点の伸びと正の関連を示した一方で、初期TOEICリーディング得点が高い学生ほど得点の伸びは小さい傾向がみられた。これらの結果は、頻度に基づく体系的な語彙学習が、大学英語教育におけるTOEICリーディングの成績向上の一要因となる可能性を示唆している。

Vocabulary knowledge plays a central role in second language development and is widely recognized as a key factor in both comprehension and communication (Schmitt et al., 2015). Despite this recognized importance, less is known about how structured vocabulary development within coordinated institutional programs relates to measurable reading outcomes over time. For Japanese university students studying English, this issue is particularly salient, as learners often differ considerably in their knowledge of high-frequency vocabulary. Despite assumptions that many core words are acquired during secondary education, university learners vary substantially in both their coverage of high-frequency vocabulary and the stability of that knowledge (Kitano & Chiba, 2020; Wadden et al., 2018). Within Japanese university EFL settings, such variability highlights the potential importance of structured vocabulary development for supporting English language growth.

The present study was conducted at a small public university, Shōzan University (a pseudonym), in central Japan offering a two-year English Language Program (ELP). Within this curriculum, the New General Service List (NGSL; Browne, Culligan, & Phillips, n.d.) serves as a core component of Year 1 instruction, with students progressing through frequency-based vocabulary levels. The NGSL's 2,800 headwords provide broad coverage of general English, making it a practical framework for structured vocabulary development in university EFL contexts.

Because TOEIC Listening and Reading (TOEIC L&R) tests are widely used in Japanese university contexts, longitudinal score data are sometimes available to track language development over time. However, improvement in skill-specific TOEIC scores is generally expected to be gradual, with prior research suggesting gains of only around 20–45 points over a two-year period (Nishitani, 2022; Wei & Low, 2017). These findings suggest that substantial increases in TOEIC Reading performance may be difficult to achieve in typical university EFL contexts, raising questions about the factors associated with stronger gains.

While the relationship between vocabulary knowledge and reading comprehension is well established (Nation, 2006; Qian & Lin, 2019; Zhang & Zhang, 2020), relatively few studies have examined how structured, frequency-based vocabulary development—such as progression through the NGSL—relates to standardized test outcomes such as TOEIC Reading. It remains unclear to what extent advancement across NGSL levels, as opposed to initial proficiency or ongoing test performance, is associated with reading gains over time in coordinated university programs. Clarifying this relationship may help institutions better understand the role of structured vocabulary development in supporting measurable English learning outcomes over time.

The purpose of this study is to examine the extent to which NGSL development, operationalized as level progression, is associated with TOEIC Reading performance over one and two years of study within a coordinated university English program. To address this question, correlational and regression analyses were conducted to explore the relationships among initial proficiency, vocabulary development, and reading gains over time. The emphasis is on providing an interpretable account of how these variables are related within an institutional context, contributing to a clearer understanding of the relationship between vocabulary development and reading gains over time.

## Literature Review

Vocabulary knowledge is among the strongest predictors of both listening and reading comprehension in a second language. For example, across 58 studies, Zhang and Zhang (2020) found that lexical knowledge accounted for roughly 45% of the variance in L2 reading outcomes. This strong linear relationship helps researchers to estimate the vocabulary size needed for comprehension at different success rates in both modalities (Hu & Nation, 2000; Qian & Lin, 2019).

Especially for reading, several researchers have proposed various thresholds depending on the level of comprehension expected or required. Estimates range from approximately 4,000–5,000-word families for 95% coverage (Laufer & Ravenhorst-Kalovski, 2010) to 8,000–9,000-word families for 98% coverage of general texts (Nation, 2006). For many Japanese university learners, vocabulary knowledge appears to remain below levels associated with minimally assisted reading of authentic texts. This lexical gap may limit learners' ability to engage with texts independently at either minimal or adequate comprehension levels, making systematic vocabulary instruction an important pedagogical concern. One practical response to this challenge has been the use of frequency-based word lists, which provide learners and educators with structured lexical

targets. Among such approaches, the NGSL has emerged as one widely used framework for organizing high-frequency vocabulary learning in university EFL contexts. As a frequency-based framework used within the present program, the NGSL provides a useful point of reference for examining relationships between vocabulary development and reading outcomes over time.

In addition to vocabulary size, depth of lexical knowledge has also been shown to play an important role in reading comprehension, particularly in processes such as lexical inferencing and the integration of word meaning in context (e.g., Qian, 2005). Recent research suggests that vocabulary knowledge operates as a multidimensional construct, with different aspects contributing in different ways to comprehension. For example, while some studies have found that receptive vocabulary depth may be more strongly associated with reading proficiency than vocabulary breadth (Tong et al., 2023), others have reported that vocabulary breadth remains the primary predictor of reading performance, with depth and fluency playing more limited roles (Zhang, 2025). While the present study focuses on vocabulary development as reflected in frequency-based level progression, it is likely that multiple dimensions of vocabulary knowledge contribute to reading performance.

A growing body of research has also examined the effects of vocabulary learning and instruction on reading development. Studies have shown that gains in lexical knowledge—whether through explicit instruction, extensive reading, or combined approaches—are associated with improvements in reading comprehension (e.g., Nation, 2001; Webb & Nation, 2017). Vocabulary learning contributes not only to increased word recognition, but also to faster processing and more efficient integration of meaning during reading, factors that may be particularly relevant in timed reading contexts such as standardized tests. However, much of this research has focused on short-term interventions or experimental conditions, and less is known about how cumulative vocabulary development within structured university programs relates to standardized reading outcomes over longer periods of time.

One approach to supporting such development is the use of frequency-based word lists. One such tool is the NGSL (Browne, Culligan, & Phillips, n.d.), which was developed to replace West's (1953) General Service List. The NGSL, comprising 2,800 headwords, was compiled from a 273-million-word contemporary corpus. Beyond enabling programs to stage instruction by level, frequency-based lists allow learners to achieve substantial text coverage with relatively few words (approximately 92% for general English texts). While the relationship between vocabulary knowledge and reading comprehension is well established, the impact of NGSL development on TOEIC Reading performance

remains less clear. In particular, it is unclear which aspects of learner progress—initial proficiency, vocabulary growth, or prior TOEIC performance—are most closely associated with later outcomes. To address this gap, the present study focused on the following research question: To what extent is NGSL development (as reflected in level progression) associated with TOEIC Reading performance over one and two years of study?

## Methodology

### Context and Participants

The case study was conducted at a small regional prefectural university in Nagano Prefecture, Japan. The institution's two-year ELP is required for all first- and second-year students in the Faculty of Global Management. Students complete four 100-minute English classes per week: two in the Accuracy Stream (form-focused instruction) and two in the Fluency Stream (communicative competence). In Year 2, Quarter 2, all students participate in a required short-term (2–4-week) study abroad program. In Year 1, the Accuracy Stream emphasizes foundational language knowledge, while the Fluency Stream focuses on communicative development; in Year 2, both streams shift toward more advanced and applied language use. NGSL mastery constituted 10% of the Year 1 Fluency Stream grade, supporting structured vocabulary development alongside communicative activities.

The participants comprised a single cohort of management majors ( $N = 174$ ; approximately 60% female) who entered the university in April 2023 and completed the ELP in January 2025. Approximately two-thirds entered through general entrance examinations. All participants consented to the use of anonymized placement and progress data for research and program evaluation in accordance with institutional guidelines. All data were de-identified prior to analysis, and because the study analyzed routine educational data collected during normal instruction, additional ethics review was not required under institutional policy (Shōzan University, 2025). In-house analyses of initial Computerized Assessment System for English Communication (CASEC) and TOEIC L&R scores suggested that this cohort was broadly comparable in proficiency and overall TOEIC score improvement to adjacent cohorts. Summary statistics from these institutional reports are available from the corresponding author upon reasonable request.

The program's vocabulary component followed a stepwise mastery system based on the NGSL, divided into five in-house frequency bands (Level 1: 1–700; Level 2: 701–1400; Level 3: 1401–1900; Level 4: 1901–2400; Level 5: 2401–2801). Students engaged in explicit vocabulary study both in and out of class through instructor-guided materials and assigned self-study. Following an initial placement check using a modified NGSL test (Stoeckel & Bennett, 2015), students were assigned to one of four levels covering the first 2,400 high-frequency headwords. Each quarter concluded with a level-specific mastery test comprising receptive (multiple-choice meaning recognition) and productive (target-word recall) sections, with advancement requiring a minimum score of 80% (Richard, 2018). Students who completed all five levels progressed to the New Academic Word List (NAWL) in subsequent study.

### Measures and Variables

The CASEC, a computer-adaptive test, was administered during orientation (April 2023) as a measure of general English proficiency. Because CASEC sub-scores are not independently validated, only total scores were used in the present study (The Japan Institute for Educational Measurement, Inc., n.d.). Table 1 summarizes all study variables and measurement points.

At the start of Year 1, students completed an NGSL placement check and were assigned to one of four levels within the program's five-band NGSL system. Vocabulary progression was assessed through quarterly mastery tests. To examine Year 1 vocabulary development, an NGSL/NAWL level-gain variable ( $T2-T1$ ) was calculated. Because students completed level-specific mastery tests targeting different lexical ranges, raw vocabulary scores were not directly comparable across time points. Therefore, level gain was used as an ordinal indicator of progression through the program's staged vocabulary system rather than as a directly comparable measure of vocabulary size increase.

TOEIC L&R Institutional Program tests were administered four times: April 2023 ( $T1$ ), January 2024 ( $T2$ ), November 2024 ( $T3$ ), and January 2025 ( $T4$ ). Only Reading scores were analyzed because the study focused specifically on the relationship between vocabulary development and reading performance, a relationship more directly established in prior research (Qian & Lin, 2019). Reading gain variables (e.g.,  $T2-T1$  and  $T4-T1$ ) were also calculated to examine changes in performance over time.

**Table 1**  
*Study Variables and Measurement Points*

| Variable                | Timing     | Description                              |
|-------------------------|------------|------------------------------------------|
| CASEC                   | April 2023 | General English proficiency (placement)  |
| Reading (T1)            | April 2023 | Baseline TOEIC Reading score             |
| NGSL Level (T1)         | April 2023 | Initial NGSL level                       |
| NGSL Level (T2)         | Jan 2024   | End-of-Year 1 NGSL/NAWL vocabulary level |
| NGSL Level Gain (T2-T1) | -          | Year 1 NGSL/NAWL progression             |
| Reading (T2)            | Jan 2024   | End-of-Year 1 TOEIC Reading score        |
| Reading (T3)*           | Nov 2024   | Late-Year 2 TOEIC Reading score          |
| Reading (T4)            | Jan 2025   | Final TOEIC Reading score                |
| Reading Gain (T2-T1)    | -          | Year 1 TOEIC Reading score change        |
| Reading Gain (T4-T1)    | -          | Overall TOEIC Reading score change       |

*Note.* TOEIC Reading (T3) is included for descriptive purposes but was not used as a variable in the analyses.

### Data Analysis

To address the research question, Pearson correlation analyses were conducted to examine the relationships between NGSL/NAWL level gain (T2-T1) and TOEIC Reading gains over one year (T2-T1) and two years (T4-T1). Additional correlations between NGSL/NAWL levels and TOEIC Reading scores at corresponding time points were conducted to provide contextual information about baseline and concurrent relationships.

To examine whether vocabulary development predicted Reading gains after controlling for baseline proficiency, multiple regression analyses were conducted using TOEIC Reading gain (T2-T1 and T4-T1) as outcome variables and CASEC, initial TOEIC Reading (T1), initial NGSL level (T1), and NGSL/NAWL level gain (T2-T1) as predictors. All analyses were conducted using JASP 0.19.3. Prior to analysis, variables were screened for missing data, normality, and multicollinearity (i.e., all VIFs < 2). Three cases were excluded because of incomplete longitudinal data resulting from program withdrawal. No transformations were required.

### Results

Table 2 summarizes the variables examined in this paper. Students entered the program with an average CASEC score of approximately 558, corresponding to the CEFR A2 range (The Japan Institute for Educational Measurement, Inc., n.d.). Mean NGSL levels increased from 2.51 to 4.31 during Year 1, corresponding to an average gain of 1.80 levels. Mean TOEIC Reading scores increased from 182 at entry to 268 at T4, an overall gain of approximately 86 points. Score dispersion also widened over time (i.e., SD = 58–73).

**Table 2**  
*Descriptive Statistics for Predictor and Dependent Variables*

| Variable                | M      | SD    |
|-------------------------|--------|-------|
| CASEC (T1)              | 557.94 | 77.48 |
| Reading (T1)            | 181.74 | 57.73 |
| NGSL Level (T1)         | 2.51   | 1.10  |
| NGSL Level (T2)         | 4.31   | 1.19  |
| NGSL Level Gain (T2-T1) | 1.80   | 0.99  |
| Reading (T2)            | 223.41 | 69.50 |
| Reading (T3)            | 257.20 | 69.49 |
| Reading (T4)            | 267.96 | 73.17 |
| Reading Gain (T2-T1)    | 42.72  | 58.42 |
| Reading Gain (T4-T1)    | 84.25  | 64.95 |

*Note.* Sample sizes varied across variables due to missing data ( $N = 159-171$ ).

To contextualize score gains, the proportion of students exceeding two commonly used thresholds for measurement error, the standard error of difference (SEdiff) and reliable change index (RCI), were also examined. Although reporting TOEIC score gains relative to such thresholds remains relatively uncommon in university-based TOEIC reporting, these indices may provide a more cautious interpretation of score change by distinguishing observed gains from those likely to reflect measurement error. For Year 1 Reading gains (T2-T1), 59.9% of students exceeded the SEdiff threshold, while 34.6% exceeded the more conservative RCI threshold. For overall gains (T4-T1), 83.0%

exceeded the SEDiff threshold and 59.1% exceeded the RCI threshold. Taken together, these findings suggest that improvement beyond measurement variability was common over the two-year period.

As shown in Table 3, both NGSL level gain (T2–T1) and final NGSL level (T2) were positively associated with TOEIC Reading gains over one year and two years, although correlations were modest in magnitude. Reading gains at T2–T1 and T4–T1 were strongly associated, suggesting continuity in growth over time. Initial Reading scores (T1) were negatively associated with subsequent gains, indicating smaller improvements among initially higher-proficiency students.

**Table 3**  
*Correlations among TOEIC Reading Gains, Initial Reading Scores, and Final NGSL Level*

| Variable                | Reading Gain (T4-T1) | Reading Gain (T2-T1) | Reading (T1) | CASEC   | NGSL Level Gain (T2-T1) |
|-------------------------|----------------------|----------------------|--------------|---------|-------------------------|
| Reading Gain (T2-T1)    | .581***              | —                    |              |         |                         |
| Reading (T1)            | -.304***             | -.275***             | —            |         |                         |
| CASEC (Placement)       | .097                 | .172*                | .509***      | —       |                         |
| NGSL Level Gain (T2-T1) | .266***              | .223**               | .105         | .148    | —                       |
| NGSL Level (T2)         | .238**               | .226**               | .574***      | .600*** | .509***                 |

Note. \* $p < .05$ , \*\* $p < .01$ , \*\*\*  $p < .001$ .

A multiple regression analysis was also conducted to determine whether NGSL level gain (T2–T1) predicted Reading gain (T4–T1) after controlling for baseline proficiency. The overall model was significant,  $F(4,151) = 17.30$ ,  $p < .001$ , explaining 31.4% of the variance in Reading gain ( $R^2 = .314$ ). Initial Reading scores (T1) were a significant negative predictor ( $\beta = -.624$ ,  $p < .001$ ), whereas NGSL level gain ( $\beta = .465$ ,  $p < .001$ ) and initial NGSL level ( $\beta = .460$ ,  $p < .001$ ) were significant positive predictors. A parallel model predicting Year 1 Reading gain (T2–T1) yielded a similar pattern of results ( $R^2 = .326$ ). In addition, NGSL level gain (T2–T1) ( $\beta = .405$ ,  $p < .001$ ) and initial NGSL level (T1) ( $\beta = .415$ ,  $p < .001$ ) were significant positive predictors.

## Discussion

The present study focused on the extent to which NGSL development, operationalized as level progression, was associated with TOEIC Reading performance over one and two years of study. The results indicate a consistent and moderate association across both timeframes. NGSL/NAWL level gain (T2–T1) was a significant positive predictor of Reading gains at both T2–T1 and T4–T1, suggesting that students who progressed further through the NGSL levels tended to show stronger Reading gains. In addition, initial NGSL level (T1) was also a significant positive predictor, indicating that students who began with stronger lexical foundations experienced greater gains over time. At the same time, initial Reading scores (T1) were negatively associated with subsequent gains, possibly reflecting ceiling effects. Taken together, these findings suggest that NGSL development is meaningfully associated with TOEIC Reading performance, with both starting lexical level and subsequent vocabulary growth contributing to reading outcomes.

The present findings should also be interpreted in relation to prior TOEIC research reporting generally modest score gains over time (Wei & Low, 2017) and studies conducted in Japanese university contexts (Nishitani, 2022). The present cohort demonstrated substantially larger average Reading gains over two years than those reported in prior research. While multiple factors likely contributed to this pattern, the findings suggest that structured vocabulary progression may represent one factor associated with stronger Reading growth within a coordinated university English program.

Across both one-year and two-year regression models, vocabulary development remained positively associated with Reading gains even after controlling for baseline proficiency. This pattern is consistent with prior research emphasizing the importance of lexical knowledge for reading comprehension (Nation, 2006; Qian & Lin, 2019; Zhang & Zhang, 2020). Although causal claims cannot be made, the findings suggest that structured vocabulary progression may be one important component associated with stronger reading development in university EFL contexts.

The findings are broadly consistent with the design of the ELP. By embedding NGSL study into the Year-1 Fluency stream and linking advancement to assessment, the curriculum provided regular reinforcement and tangible milestones. The clear upward movement of most students from Levels 1–3 to Levels 4–5 and NAWL indicates that a frequency-based, cumulative system was associated with measurable progress among these non-English majors. Because the initial NGSL levels check assessed only the first

three NGSL levels, no students began at Level 5, which may have introduced a slight ceiling effect in the baseline data; nevertheless, subsequent advancement patterns show that the program successfully differentiated learner growth over time.

In practical terms, the ELP's step-by-step vocabulary sequence functions as a foundation for developing reading competence. Students who advanced systematically through the NGSL bands achieved stronger gains in reading proficiency, and consistent effort and structured exposure appeared to compensate for weaker starting points. In a coordinated program such as that at Shōzan University, what students learn in Year 1 continues to matter deeply in Year 2. In short, the findings suggest that Year 1 vocabulary development may continue to relate to later reading performance. This confirms the program's guiding assumption that incremental lexical development is essential for broader language competencies.

For educators, several implications emerge from these findings. First, learners may benefit from explicit and cumulative tracking of vocabulary development. Making progress through frequency-based bands visible to students may help support sustained engagement with lexical learning while emphasizing mastery over memorization. Second, regular formative assessment appears valuable. Because vocabulary development and Reading gains were positively associated across both one-year and two-year periods, staged assessment may provide students and instructors with useful feedback on progress over time. Finally, the present study illustrates how routinely collected institutional data can be used to explore program outcomes. In contexts where standardized assessments are already administered, relatively simple analyses may offer a practical entry point for evidence-informed curriculum reflection. Although the present study was conducted within a coordinated curriculum, the underlying approach to vocabulary tracking may be adaptable to less coordinated contexts. Frequency-based vocabulary targets, staged progression, and periodic assessment can be implemented at the classroom level, even when instructors work independently of broader program structures.

## Conclusion

The present study illustrated the extent to which NGSL development, operationalized as level progression, was associated with TOEIC Reading gains over one and two years within a university ELP. The findings indicated a consistent positive association between NGSL/NAWL development and Reading gains across both timeframes. Students who progressed further through the program's staged vocabulary system tended to show stronger gains in TOEIC Reading performance, even after controlling for baseline

proficiency. Although causal claims cannot be made, the results are consistent with prior research, highlighting the importance of vocabulary knowledge for reading development.

These findings suggest that structured, frequency-based vocabulary learning may represent one meaningful component associated with reading development in university EFL contexts. While the present study was conducted within a coordinated curriculum, elements of the approach—such as staged vocabulary targets and periodic assessment—may be adaptable to less coordinated teaching settings. At the same time, the study's single-cohort design limits generalizability, and vocabulary development may also reflect broader learner characteristics, including motivation and engagement, which were not directly measured.

Future research should examine whether similar patterns emerge across additional cohorts and institutional contexts, while also exploring how vocabulary development correlates with other aspects of language learning over time. Overall, the findings suggest that sustained vocabulary development may play an important role in supporting Reading growth among non-English-major university learners in Japan.

## Bio Data

**Jean-Pierre J. Richard**, EdD, is an associate professor in the Faculty of Global Management at Shōzan University. His research focuses on vocabulary acquisition, individual differences, and assessment. He is currently developing NGSL vocabulary learning materials for students with lower proficiency levels, is leading a collaborative textbook project for academic discussion, and is investigating how TOEIC results are reported, interpreted, and used in Japanese higher education. richard.jean-pierre@u-nagano.ac.jp

## References

- Browne, C., Culligan, B., & Phillips, J. (n.d.). *New General Service List Project*. <http://www.newgeneralservicelist.com>
- Hu, M., & Nation, P. (2000). Unknown vocabulary density and reading comprehension. *Reading in a Foreign Language*, 13(1), 403–430. <https://nflrc.hawaii.edu/rfl/item/43>
- Kitano, M. L. & Chiba, K. (2020). Analyzing gaps of knowledge in high frequency English words by university freshmen in Japan. *Bunkyo University Faculty of International Studies*, 31(1), 15–23. <https://bunkyo.repo.nii.ac.jp/records/7928>

- Laufer, B., & Ravenhorst-Kalovski, G. C. (2010). Lexical threshold revisited: Lexical text coverage, learners' vocabulary size and reading comprehension. *Reading in a Foreign Language*, 22(1), 15–30. <https://doi.org/10.64152/10125/66648>
- Nation, I. (2006). How large a vocabulary is needed for reading and listening? *The Canadian Modern Language Review*, 63(1), 59–82. <https://doi.org/10.3138/cmlr.63.1.59>
- Nishitani, A. (2022). Curricular innovation impact analysis: Parallel process growth curve models. In S. J. Ross & M. C. Masters (Eds.), *Longitudinal studies of second language learning* (pp. 150–170). Routledge. <https://doi.org/10.4324/9781003087939-9>
- Qian, D. D. (2005). Demystifying lexical inferencing: The role of aspects of vocabulary knowledge. *TESL Canada Journal*, 22(2), 34–54.
- Qian, D. D., & Lin, L. H. F. (2019). The relationship between vocabulary knowledge and language proficiency. In S. Webb (Ed.), *The Routledge handbook of vocabulary studies*, (pp. 66–80). Routledge. <https://doi.org/10.4324/9780429291586>
- Richard, J-P. J. (2018). Developing written productive items for an institution-level NGSL assessment program. *Vocabulary Education & Research Bulletin*, 7(2), 14–20.
- Schmitt, N., Cobb, T., Horst, M., & Schmitt, D. (2015). How much vocabulary is needed to use English? Replication of van Zeeland & Schmitt (2012), Nation (2006), and Cobb (2007). *Language Teaching*, 50(2), 212–226. <https://doi.org/10.1017/S0261444815000075>
- Shōzan University. (2025). *English program for global mobility: 2025 student handbook*. [Institutional document].
- Stoeckel, T., & Bennett, P. (2015). A test of the new general service list. *Vocabulary Learning and Instruction*, 4(1), 1–8. <https://doi.org/10.7820/vli.v04.1.stoeckel.bennett>
- The Japan Institute for Educational Measurement, Inc. (n.d.). CASEC: 他英語試験との関係 [CASEC: Relationship with other English tests]. <https://casec.evidus.com/materials/>
- Tong, Y., Hasim, Z. & Abdul Halim, H. A. (2023). The relationship between L2 vocabulary knowledge and reading proficiency: The moderating effects of vocabulary fluency. *Humanities and Social Sciences Communications*, 10. <https://doi.org/10.1057/s41599-023-02058-2>
- Wadden, P., Browne, C., & Nation, P. (2018). Teaching and learning vocabulary in the Japanese university. In P. Wadden & C. C. Hale (Eds.), *Teaching English at Japanese universities* (pp. 84–96). Routledge. <https://doi.org/10.4324/9781315147239>
- Webb, S. & Nation, P. (2017). *How vocabulary is learned*. Oxford University Press.
- Wei, Y., & Low, A. (2017). Monitoring score change patterns to support TOEIC listening and reading test quality. *ETS Research Reports Series*, 2017(1), 1–20. <https://onlinelibrary.wiley.com/doi/10.1002/ets2.12186>
- West, M. (1953). *A general service list of English words*. Longman.
- Zhang, S. & Zhang, X. (2020). 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>
- Zhang, X. (2025). Vocabulary knowledge towards L2 reading and listening performance. *Reading in a Foreign Language*, 37(2), 109–126. <https://hdl.handle.net/10125/67500>