TESL-EJ Teaching English as a Second or Foreign Language
*The Electronic Journal for English as a Second Language*

# The Role of Vocabulary Learning Strategies in Longitudinal Vocabulary Growth

**August 2026 – Volume 30, Number 2**
**https://doi.org/10.55593/ej.30118a6**

**Masaki Akase**
National Institute of Technology, Nagano College, Japan
<m_akase@nagano-nct.ac.jp>

## Abstract

This study examines the role of vocabulary learning strategies (VLS) in vocabulary size (VS) and vocabulary growth among science, technology, engineering, and mathematics (STEM)-oriented upper secondary students in Japan. Using a three-year longitudinal mixed-methods design, 189 participants completed three forms of the Vocabulary Size Test (VST), a fourth form composed of items from VSTs 1–3, and a VLS questionnaire administered at four time points. Total VS was estimated using a Rasch-based method to ensure comparability across test forms. Cluster analysis identified three strategy-use groups (low, medium, high), and a two-way mixed ANOVA revealed significant main effects of time and strategy cluster, as well as their interaction, indicating that more frequent strategy use was modestly but consistently associated with vocabulary growth, with small but cumulative effects over time. Open-ended responses further revealed differences in strategy depth and diversity across clusters. Integrating quantitative and qualitative findings, this study identifies patterns of vocabulary development among students at Japanese Colleges of Technology (KOSEN), underscores the importance of sustained strategy engagement, and offers implications for vocabulary assessment and longitudinal monitoring in diverse EFL contexts, including technical and higher education settings.

***Keywords:*** vocabulary learning strategies, vocabulary size, Japanese EFL learners, longitudinal study, Rasch measurement

## Background to the Study

Vocabulary is central to foreign language learning; as Wilkins (1972, p. 111) famously observed, "without grammar very little can be conveyed, without vocabulary nothing can be conveyed." Adequate vocabulary supports both comprehension and expression, providing a foundation for all four language skills. Reflecting this importance, the Japanese Ministry of Education, Culture, Sports, Science and Technology (MEXT) increased the target number of English words taught from approximately 3,000 to 4,000–5,000 across elementary to high school levels (MEXT, 2018). Nevertheless, a nationwide survey reported that 71.8% of first-year high school students found vocabulary learning difficult (Benesse, 2019), suggesting that learners' actual acquisition may fall short of policy expectations. Moreover, Nation (2006) estimated that roughly 8,000 word families are required to reach 98% lexical coverage for adequate comprehension, indicating that Japanese EFL learners face a substantial and cumulative vocabulary burden.

Given that vocabulary knowledge underpins communicative competence, an important question arises regarding which VLS are associated with vocabulary growth. Fostering the use of VLS may support learner autonomy and engagement (Nation, 2022). Understanding which types of strategies are most effective, and how their use evolves over time, is therefore critical for providing both theoretical insights and practical guidance in EFL instruction. The present study investigates these questions in the context of Japanese EFL learners enrolled in a five-year STEM-oriented educational program at KOSEN, aiming to clarify the relationship between VLS use and VS growth.

## Literature Review

Vocabulary knowledge is commonly conceptualized in terms of size and depth (Schmitt, 2014). VS, or breadth, refers to the quantitative dimension of lexical knowledge—that is, the number of words a learner knows at a basic level, including knowledge of word forms and primary meanings (Anderson & Freebody, 1981; Nation, 2022; Qian, 1999; Read, 2000). In contrast, vocabulary depth represents the qualitative aspects of word knowledge, including how lexical items are organized and interconnected within the mental lexicon (Read, 2000; Schmitt, 2008; Stæhr, 2009).

Previous research has emphasized the importance of VS because of its central role in language use and overall language proficiency (Laufer et al., 2004; Meara, 2002; Nation, 2013; Schmitt, 2010; Schmitt & Schmitt, 2014). VS is especially informative for learners with limited lexical repertoires (Meara, 1996) and is often considered a more representative indicator of overall vocabulary knowledge than depth measures (Read, 2000). In addition, VS shows strong associations with general language proficiency and performance across the four skills—reading, listening, writing, and speaking (Milton & Hopwood, 2023). It also contributes to reading and listening comprehension (Webb & Paribakht, 2015) and facilitates incidental vocabulary learning (Teng & Uchihara, 2024).

Considering its theoretical and pedagogical importance, VS is commonly operationalized and measured through simple recognition tests of word forms and meanings. Such assessments are widely used as indicators of learners' overall vocabulary proficiency across the four language skills. For beginning EFL learners, including junior and senior high school students, vocabulary breadth—that is, basic knowledge of many words—may be more immediately relevant than deeper lexical knowledge. Although these measures capture only surface-level aspects of vocabulary knowledge, they offer a widely accepted and practical approach to evaluating vocabulary breadth and tracking developmental change over time, particularly in classroom-based longitudinal research.

### Measuring Vocabulary Size

VS is a robust predictor of reading, writing, general language proficiency, and academic achievement in EFL contexts (Laufer & Goldstein, 2004). Several instruments, including the VST (Nation & Beglar, 2007) and the Updated Vocabulary Levels Test (Webb et al., 2017), assess learners' knowledge of word forms and their associated meanings, often in bilingual formats. While test formats are similar, they differ in corpus selection and sampling, typically using stratified-random sampling by frequency level (Schmitt, 2010). VS is estimated from the proportion of known words per level, multiplied by 1000, following the conventional raw-score approach, although Rasch-based scaling has increasingly been adopted for improved measurement precision.

Recent research has emphasized rigorous validity assessment for VST development (Schmitt et al., 2019). Rasch analysis (Rasch, 1960), a one-parameter item response theory model, has been widely applied to evaluate unidimensionality—that is, whether responses can be

explained by a single underlying latent trait—as well as item fit, reliability, and linking of different test versions (Henning et al., 1985). Rasch difficulty estimates and fit statistics can be used to evaluate whether individual items contribute appropriately to the reliability of the test as a whole (Bond & Fox, 2001). Rasch analysis can also be used to link different versions of a test to a single scale, so long as the versions have some items in common (Bond et al., 2020).

Several studies have validated the VST for Japanese EFL learners (Beglar, 2010; Hamada et al., 2021; Kasahara, 2006; Koizumi & Mochizuki, 2011). Kasahara (2006) revised Mochizuki's (1998) VST using FACETS Rasch software, replacing misfitting items with items from the JACET 8000 Basic Words list (JACET, 2003) to enhance unidimensionality, reliability, and validity. Beglar (2010) validated the VST created by Nation and Beglar (2007) with a large group of learners, confirming good Rasch fit, unidimensionality, reliability, and invariance. Hamada et al. (2021) developed a VST based on the New JACET List, validated it with a three-parameter logistic IRT model, and demonstrated that it predicted reading and listening proficiency. Koizumi and Mochizuki (2011) also applied the Rasch model to Mochizuki's VST and found evidence of measurement consistency, including few misfitting items, increasing item difficulty across frequency levels, and a strong correlation with TOEIC® L&R scores.

The VST can also track changes in learners' VS over time using conventional raw-score methods (e.g., Akase, 2021; Katagiri, 2009). Building on this line of research, the present study adopts the equated VST forms developed by Akase (2022), which linked three original versions created by Aizawa and Mochizuki (2010) and a newly created fourth form through Rasch analysis, thereby providing reliable longitudinal measures of VS. Based on Rasch person ability estimates, the estimated number of words known in each frequency band was calculated as the mean probability of success for that level multiplied by 1000. Total VS was then obtained by summing the estimated numbers of words known across all frequency levels, yielding Rasch-based VS estimates for subsequent analyses.

**Vocabulary Learning Strategies**

Since the mid-1980s, second language acquisition research has highlighted the pivotal role of VLS in vocabulary acquisition and overall language proficiency. VLS are broadly defined as "actions that learners take to help themselves understand and remember vocabulary" (Cameron, 2001, p. 92). These include memorization, repetition, association, use of the keyword method—which links the meaning of an L2 word with an L1 word that has a similar spoken form and has been shown to support durable long-term vocabulary learning (Webb & Nation, 2017)—as well as contextualization and dictionary use (Nyikos & Fan, 2007). Systematic investigation of VLS effects on VS has been enabled by widespread VST implementation (Gu & Johnson, 1996; Kojic-Sabo & Lightbown, 1999). While some strategies, such as visual repetition and rote memorization, have been found to correlate negatively with VS, most show positive associations.

Research examining the relationship between VLS and VS has identified several consistent patterns. First, learners with higher proficiency and larger VS tend to use a wider variety of strategies more frequently than less successful learners (e.g., Fan, 2003; Gu, 2010; Lawson & Hogben, 1996). Second, as learners gain proficiency, their strategy use gradually shifts from simple strategies, such as rote repetition, to more complex strategies, including imagery, inferencing, association, and organization (e.g., Horino & Ichikawa, 1997; Lai, 2016; Schmitt, 1997; Zhang & Lu, 2015). Third, the appropriate, structured, and balanced use of strategies is associated with more successful vocabulary learning and higher proficiency (e.g., Akase, 2021; Mizumoto & Takeuchi, 2009a; Sanaoui, 1995). Fourth, metacognitive strategies and cognitive strategies—particularly memory strategies—play a central role in effective vocabulary learning, and explicit strategy instruction has been demonstrated to enhance learners' VS (e.g., Atay &

Ozbulgan, 2007; Mizumoto & Takeuchi, 2009b; Rasekh & Ranjbary, 2003). Finally, individual difference variables, including gender, age, academic discipline, and learning context, influence both the choice and frequency of VLS use (e.g., Catalán, 2003; Nakamura, 2002; Tseng & Schmitt, 2008).

Despite a substantial body of research on VLS, most studies rely on cross-sectional designs. This leaves unanswered questions about how VLS use develops over time and how such development influences VS growth. As Boers (2024) argues, adopting a longitudinal perspective is essential for understanding not only how learning strategies are acquired but also how they can be effectively evaluated. Considering the crucial role of VS in communicative competence, examining the dynamic interplay between VS and VLS longitudinally can provide important theoretical and pedagogical insights (Gu, 2019).

To address this gap, the present study investigates patterns of VLS use and their relationship with VS development through both cross-sectional and longitudinal analyses. Given that VS provides a strong indication of learners' general linguistic progress toward communicative goals (Milton & Hopwood, 2023), the investigation is guided by four research questions (RQs):

> **RQ1:** What types of VLS do learners use, and to what extent are these strategies employed at Time 4 (the final measurement point)?
>
> **RQ2:** What patterns of VS and VS growth emerge in relation to changes in learners' VLS use over time?
>
> **RQ3:** To what extent can four parallel forms of the VST, together with a VLS questionnaire, be used to capture relationships between patterns of VLS use and VS growth?
>
> **RQ4:** Based on text mining of learners' open-ended responses, which specific VLS can be identified?

## Method

This study employed a quantitative longitudinal design to examine changes in VS and VLS over time. Data were collected using four parallel forms of the VST and a VLS questionnaire administered at multiple measurement points.

### Participants

The participants were 189 students (36 females and 153 males; aged 16–18 years) enrolled in the first three years of a five-year STEM program at a Japanese college of technology (KOSEN), a tertiary institution that admits students after junior high school and provides specialized engineering education. This period corresponds to the senior high school stage, during which learners typically experience substantial vocabulary growth within a consistent English curriculum. The researcher taught these students throughout this period, ensuring stable instructional conditions and minimizing external variation. Although the researcher also served as the instructor, testing followed standardized procedures. Prior to data collection, students were informed of the purpose of the study, that participation was voluntary, and that they could withdraw from the study at any time without penalty.

Students' English proficiency ranged from A2 to B1 on the CEFR, with a few students reaching B2. Across five engineering-related departments—mechanical engineering, electrical and electronic engineering, electronics and control engineering, computer science, and civil engineering—English was regarded as an important academic and professional skill. About half of the students advanced to higher education, while the remainder entered the workforce, making English an important component of their academic and professional development.

## Instruments

**Vocabulary Size Test**. Data were collected at four time points: Time 1 (April 2018, first year), Time 2 (April 2019, second year), Time 3 (April 2020, third year), and Time 4 (December 2020, third year). Three forms of Aizawa and Mochizuki's (2010) VST, along with an equated version, VST 4, developed by Akase (2022), were administered. These tests use Japanese definitions and are suitable for lower-intermediate EFL learners. To balance test length, difficulty, and reliability, six frequency bands (1,000–6,000 words) were included; higher levels were excluded to reduce guessing, which could otherwise overestimate VS and compromise measurement accuracy (Stewart, 2014). Each band contained 26 items, resulting in 156 items per form.

VST 4 employed a Rasch equating procedure to align the original three forms, thereby enabling reliable longitudinal measurement while minimizing practice effects. Items were carefully selected from VSTs 1–3 to cover the full difficulty range and ensure representation across each frequency band (see Appendix A in the Supplementary File). Rasch equating also allows estimation of the expected number of correct responses per frequency band by averaging item difficulties, from which total VS can be inferred. This approach ensures valid longitudinal comparisons and supports robust analyses of VS development.

**Vocabulary Learning Strategies Questionnaire.** An inventory of VLS was compiled from the literature to develop the Vocabulary Learning Strategies Questionnaire (VLSQ), drawing on instruments by Fan (2003), Gu and Johnson (1996), Horino and Ichikawa (1997), Mizumoto and Takeuchi (2009a), and Schmitt (1997). Following Dörnyei's (2005) observation that "it is difficult to distinguish the difference between engaging in an ordinary learning activity and a strategic learning activity" (p. 143), VLS in this study are defined as deliberate or goal-directed actions, techniques, or behaviors that learners employ to acquire and practice vocabulary (cognitive strategies) or to plan, monitor, and regulate their vocabulary learning (metacognitive strategies) (Fan, 2003; Oxford, 1990; Schmitt, 1997, 2000). While this conceptualization emphasizes intentional strategy use, some strategies may become automatic or habitual with repeated practice, reflecting the dynamic nature of strategy development over time (Cohen, 2011; Dörnyei, 2005; Griffiths, 2018).

Based on this framework, 33 items were developed to assess seven strategy categories: metacognitive (6), input-seeking (4), written rehearsal (5), oral rehearsal (3), social (4), organization (6), and elaboration (5). Participants rated their strategy use on a 6-point Likert scale (1 = strongly disagree, 6 = strongly agree). The questionnaire items are provided in Appendix B in the Supplementary File. Table 1 presents an overview of the quantitative instruments used in the study.

**Table 1. Summary of Quantitative Measures Used in the Study**

| Measure | Items | Possible range |
|---|---|---|
| Vocabulary size test (VST 1, 2, 3, 4) | 156 / VST | 0–6000 |
| Metacognitive strategies | 6 | |
| Input-seeking strategies | 4 | |
| Written rehearsal strategies | 5 | |
| Oral rehearsal strategies | 3 | 1–6 |
| Social strategies | 4 | |
| Organization strategies | 6 | |
| Elaboration strategies | 5 | |

The VLSQ was administered alongside the four VST forms at each measurement point. For the present analysis, only Time 4 VLS data were used to represent learners' habitual strategy use,

as strategy use tends to stabilize with maturation (Akase, 2021) and become more automatic with experience (Griffiths, 2018). By the third year, learners had accumulated substantial instructional exposure, making Time 4 a reasonable proxy for relatively stable strategy repertoires and typical strategy profiles. This approach was intended to provide a relatively stable index of learners' strategy use and to align the VLS data with the longitudinal VS analyses.

Because the analysis relies on end-point strategy data, no causal claims are made regarding earlier vocabulary growth, and associations between strategy use and vocabulary outcomes are interpreted cautiously.

**Open-Ended Survey on Vocabulary Learning.** At the end of their third year (March 2021), participants completed a voluntary open-ended online survey administered via Google Forms. Of the 189 students invited, 178 responded (36 females, 142 males) to the prompt: "What techniques or strategies do you use specifically when learning vocabulary?" The online format allowed learners to respond at their convenience, reduced transcription burden, minimized researcher influence, and facilitated anonymous data collection.

In contrast to the structured VLSQ, these open-ended responses provided self-generated descriptions of strategy use, including strategies not fully captured by the predefined questionnaire categories. The qualitative data played a complementary role in explaining the quantitative clusters and interpreting differences among strategy groups. Specifically, themes derived from the open-ended responses were used to elaborate, confirm, and nuance patterns identified statistically.

Following an explanatory sequential mixed-methods design (see Figure 1), the study first collected and analyzed longitudinal quantitative data on VLS use and VS, then used qualitative data to explain and contextualize the results (Creswell, 2015). Although collected at a single time point, the qualitative data were not intended to establish causality but to illuminate patterns observed in the longitudinal analyses. By integrating both strands at the interpretation stage, the study provides a more comprehensive and theoretically grounded understanding of learners' VLS and their association with VS development.

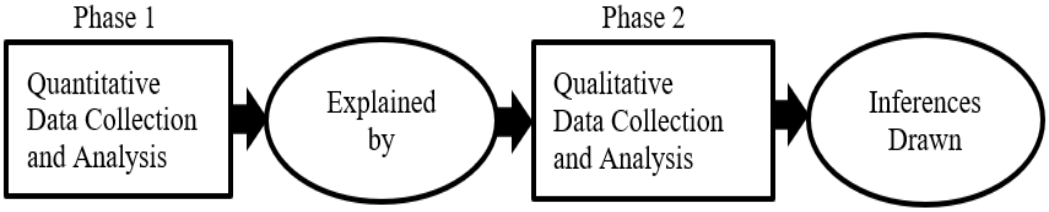


**Figure 1. Explanatory Sequential Design**
**Note. Adapted from Creswell (2015, p. 39).**

## Procedures

Data collection followed institutional approval and cooperation from course instructors across four measurement points (Time 1–4). At each point, students completed the VST and VLSQ under standardized classroom conditions. An open-ended online survey was administered at the end of the third year (March 2021) to capture learners' self-generated descriptions of vocabulary learning practices beyond the structured questionnaire.

All data were de-identified and compiled in Excel. Quantitative analyses were conducted using Winsteps 4.7 (Linacre, 2019) for Rasch calibration of the VST forms (Akase, 2022). Using the dichotomous Rasch model, item difficulty and learner ability were estimated on the same latent scale. Descriptive statistics, principal component analysis (PCA), hierarchical cluster analysis, and repeated-measures mixed ANOVA were conducted using JASP 0.12.2 (JASP Team, 2020).

These procedures yielded relatively stable estimates of VS and strategy-use profiles, which informed the qualitative phase. Cluster analysis enabled pedagogically interpretable group comparisons and examined whether different levels of strategic engagement were associated with distinct longitudinal vocabulary growth trajectories.

Following an explanatory sequential mixed-methods design (Creswell, 2015), qualitative analyses were guided by the quantitative results. Open-ended responses were linked to cluster membership and analyzed using KH Coder 3.0 (Higuchi, 2014) to identify salient keywords and recurrent themes for each strategy group. This integration allowed qualitative findings to elaborate on, qualify, and interpret patterns observed in the quantitative analyses, thereby enhancing the explanatory depth and validity of the study.

### Data analysis

The data were analyzed using quantitative statistical procedures. Descriptive statistics (means and standard deviations) were calculated, and the reliability of the questionnaire was assessed using Cronbach's alpha. Cluster analysis was conducted to identify profiles of VLS use, followed by a two-way mixed ANOVA to examine differences in VS development across the identified groups over time.

## Results

### Descriptive Statistics for the Vocabulary Size Tests

Table 2 presents descriptive statistics for the 189 participants across VSTs 1–4. Rasch-estimated VS, equated across test forms, increased steadily over time. The distributions of VS scores were approximately normal at each wave (skewness and kurtosis between −1.5 and +1.5; Tabachnick & Fidell, 2013), supporting the use of parametric analyses in subsequent longitudinal comparisons.

**Table 2. Descriptive Statistics for the VST Scores**

| VST | Time | *M* | *SD* | Skewness | Kurtosis | *k* | *α* |
|---|---|---|---|---|---|---|---|
| VST 1 | 1 | 2734.31 | 486.63 | 0.00 | 0.05 | 156 | .87 |
| VST 2 | 2 | 3035.44 | 652.33 | -0.44 | 0.29 | 156 | .93 |
| VST 3 | 3 | 3401.77 | 661.30 | 0.01 | -0.24 | 156 | .93 |
| VST 4 | 4 | 3734.34 | 745.43 | -0.24 | 0.51 | 156 | .94 |

**Note. *N* = 189; *k* = number of items; VST = Vocabulary Size Test.**

### Validation of a Vocabulary Learning Strategy Questionnaire

PCA with Oblimin rotation was conducted to examine the underlying structure of the 33-item VLSQ and to identify components for subsequent analyses. This approach was selected as an appropriate data-reduction technique for identifying coherent item groupings while allowing correlations among components, which are theoretically expected in strategy use.

An initial seven-component solution (eigenvalues > 1.0; scree plot inspection) largely yielded loadings above .40. However, one item cross-loaded, two items failed to load clearly, and the seventh component had a marginal eigenvalue (1.15) with limited interpretability. Considering these results alongside theoretical frameworks (Dörnyei & Dewaele, 2022; Gu & Johnson, 1996; Schmitt, 1997), a six-component solution was adopted. Three items (VLS-14, VLS-15, VLS-26) were removed, resulting in a final 30-item instrument. This refined structure balances empirical adequacy with theoretical coherence, providing support for the construct validity of the instrument.

The six components were interpreted and labeled as follows:

1. Metacognitive strategies: input-seeking and regulation of learning (Fan, 2003; Mizumoto & Takeuchi, 2009a; Schmitt, 1997).

2. Grouping strategies: organizational and memory-based word association strategies (Horino & Ichikawa, 1997; Schmitt, 1997).

3. Contextualizing strategies: imagery and contextual use of target vocabulary (Gu & Johnson, 1996).

4. Traditional strategies: focused practices such as flashcards, vocabulary notebooks, and written rehearsal (Mizumoto & Takeuchi, 2009a; Schmitt, 1997).

5. Social strategies: collaborative learning and peer interaction (Schmitt, 1997).

6. Oral rehearsal strategies: repeating and practicing vocabulary aloud (Schmitt, 1997).

Table 3 presents descriptive statistics and reliability indices for the six VLSQ components at Time 4. Mean scores for metacognitive and grouping strategies were around 3, indicating moderate-to-low levels of use, while traditional and social strategies were used even less frequently. In contrast, contextualizing and oral rehearsal strategies were employed more often, with mean scores of approximately 3.8. Cronbach's alpha coefficients exceeded .70 for all components except oral rehearsal strategies, which demonstrated marginal but acceptable reliability given its limited number of items.

**Table 3. Descriptive Statistics for the Six Components of VLS at Time 4**

| VLS | *k* | Items | *M* | *SD* | Skewness | Kurtosis | *α* |
|---|---|---|---|---|---|---|---|
| MET | 8 | 2, 10, 12, 20, 22, 24, 29, 33 | 3.12 | 1.00 | 0.02 | -0.43 | .83 |
| GRP | 6 | 1, 8, 11, 23, 30, 31 | 3.21 | 0.98 | -0.09 | -0.23 | .74 |
| CON | 7 | 3, 6, 7, 17, 19, 21, 27 | 3.87 | 0.86 | -0.62 | 0.88 | .75 |
| TRD | 4 | 4, 25, 28, 32 | 2.71 | 1.13 | 0.51 | -0.22 | .71 |
| SOC | 2 | 5, 16 | 2.74 | 1.35 | 0.36 | -0.77 | .78 |
| ORL | 3 | 9, 13, 18 | 3.83 | 1.13 | -0.37 | -0.25 | .65 |

**Note. *N* = 189; *k* = number of items; VLS = vocabulary learning strategies; MET = metacognitive strategies; GRP = grouping strategies; CON = contextualizing strategies; TRD = traditional strategies; SOC = social strategies; ORL = oral rehearsal strategies.**

## Cluster Analysis

A cluster analysis (Ward's method, squared Euclidean distance) of the Time 4 VLSQ data yielded three clusters that reflected overall differences in strategy-use frequency: high (Cluster 1, $n = 53$), medium (Cluster 2, $n = 101$), and low (Cluster 3, $n = 35$), rather than qualitatively distinct strategy profiles. To facilitate interpretation, the clusters were labeled high-, medium-, and low-strategy-use groups.

As shown in Figure 2, between-cluster differences were most pronounced for metacognitive, grouping, and contextualizing strategies, with Cluster 1 scoring above and Cluster 3 below the overall sample mean, whereas differences were smaller for traditional, social, and oral rehearsal strategies—particularly for traditional strategies, which showed similar levels of use in Clusters 1 and 2.

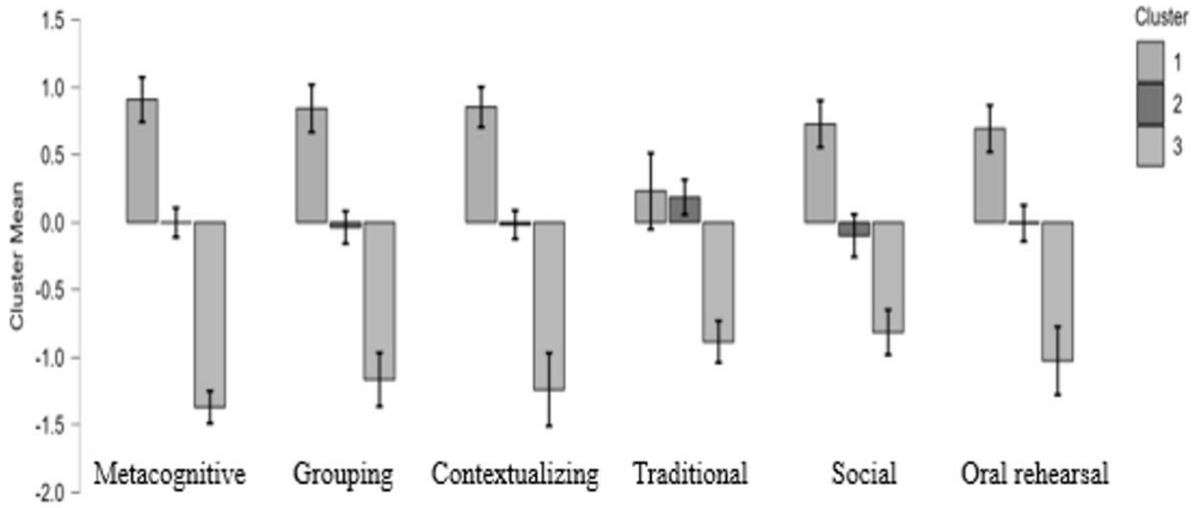


**Figure 2. Vocabulary Learning Strategy Use Cluster at Time 4**
**Note. Cluster 1 (*n* = 53), Cluster 2 (*n* = 101), and Cluster 3 (*n* = 35).**

In general, the clusters reflect differences in the overall intensity of strategy engagement, although these differences were relatively small, particularly for traditional, social, and oral rehearsal strategies. Accordingly, the cluster analysis is interpreted as providing a descriptive account of individual differences in strategy engagement rather than evidence for categorical learner types.

## Two-Way Mixed ANOVA

A two-way mixed repeated-measures ANOVA was conducted to examine differences in Rasch-estimated VS across time, with strategy clusters as a between-subjects factor and time as a within-subjects factor. Strategy clusters were treated as an index of overall strategy engagement rather than categorical learner types. Preliminary analyses indicated that the assumptions of homogeneity of variance and sphericity were met (Levene's test; Mauchly's test).

Table 4 shows significant main effects for both strategy clusters and time, suggesting that learners differed in overall VS according to their level of strategy engagement and that VS increased across the four measurement points. The significant cluster × time interaction further reflects modest differences in developmental trajectories across strategy groups.

**Table 4. ANOVA Results for Strategy Clusters and Time (Growth in VS)**

| Source | *SS* | *df* | *MS* | *F* | *p* | $\eta^2$ |
|---|---|---|---|---|---|---|
| *Between Subjects* | | | | | | |
| Strategy Clusters | 18294033.40 | 2 | 9147016.70 | 8.45 | <. 001 | .05 |
| Residuals | 20145854.50 | 186 | 1082880.94 | | | |
| *Within Subjects* | | | | | | |
| Time | 79637893.41 | 3 | 26545964.74 | 173.40 | <. 001 | .21 |
| Strategy Clusters × Time | 4440449.65 | 6 | 740074.94 | 4.83 | <. 001 | .01 |
| Residuals | 85425801.24 | 558 | 153092.83 | | | |

**Note. *N* = 189; VS = Vocabulary Size; $\eta^2$ = .01 (small), .06 (medium), .14 (large).**

Table 5 presents the mean and standard deviation of estimated VS for each strategy cluster. Figure 3 shows the longitudinal trajectories of estimated VS across strategy clusters. Follow-up analyses revealed that although VS levels were comparable at Time 1 ($p$ = .06, $\eta^2$ = .03), learners in the high-strategy-use group showed greater cumulative gains at later time points. In particular, Cluster 1 showed significantly larger increases than Cluster 3 by Time 4 ($p$ < .001, $\eta^2$ = .13).

**Table 5. Mean Estimated VS on Time (VST) of the Three Strategy Clusters**

| | Cluster 1 ($n$ = 53) | Cluster 2 ($n$ = 101) | Cluster 3 ($n$ = 35) |
|---|---|---|---|
| Time 1 (VST 1) | 2861.65 (512.00) | 2673.64 (434.43) | 2716.56 (471.71) |
| Time 2 (VST 2) | 3251.45 (689.49) | 3016.24 (628.75) | 2763.74 (637.56) |
| Time 3 (VST 3) | 3583.74 (811.34) | 3394.65 (621.80) | 3146.78 (612.26) |
| Time 4 (VST 4) | 4056.02 (748.90) | 3718.90 (606.15) | 3291.81 (710.29) |

**Note. Mean (Standard Deviation).**

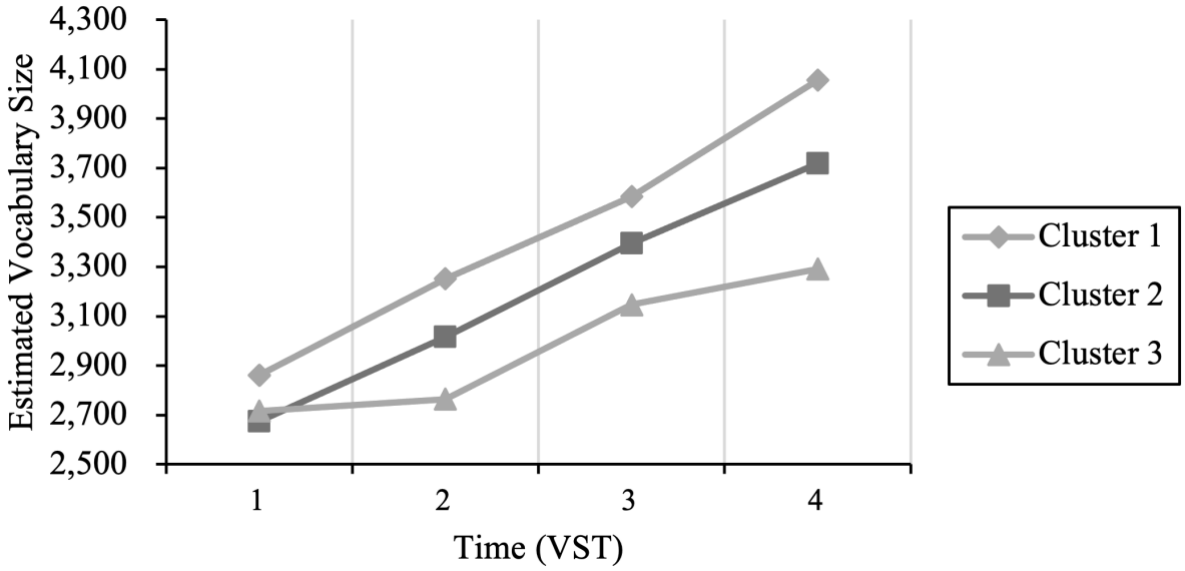


**Figure 3. Estimated Vocabulary Size of Three Strategy Clusters at Four Time Points**
**Note. Cluster 1 ($n$ = 53), Cluster 2 ($n$ = 101), and Cluster 3 ($n$ = 35).**

Overall, these findings suggest that higher levels of strategy engagement are associated with greater longitudinal vocabulary growth. Importantly, the results reflect differences in the degree of strategy engagement rather than qualitatively distinct learning pathways, and no causal claims are implied.

### Analysis of Open-Ended Survey Responses

After data cleaning procedures—including the removal of blank responses, correction of typographical errors and coding inconsistencies—the open-ended responses were analyzed using KH Coder 3.0 (Higuchi, 2014). A total of 3,436 word tokens were identified across 293 sentences in response to the open-ended question on VLS. All analyses were conducted on the original Japanese responses to preserve semantic accuracy, after which representative words and excerpts were translated into English for reporting purposes.

Following the exclusion of duplicated entries and function words, 1,565 content words were retained for analysis. Consistent with the explanatory sequential mixed-methods design, the qualitative data were analyzed in relation to the quantitative strategy clusters, allowing characteristic lexical patterns to be identified for each group.

Table 6 presents the top 10 distinctive words for each VLS cluster. These keywords highlight differences in learners' self-reported strategy behaviors and provide interpretive insight into the quantitative cluster profiles. In this way, the qualitative findings serve not merely as illustrative examples but as an interpretive layer that contextualizes and elaborates the statistical patterns observed in both the cluster and longitudinal analyses.

**Table 6. List of Distinctive Words Extracted from Each Cluster**

| No. | Cluster 1 ($n$ = 51) | | | Cluster 2 ($n$ = 97) | | | Cluster 3 ($n$ = 30) | | |
|---|---|---|---|---|---|---|---|---|---|
| | Japanese | English | Jaccard | Japanese | English | Jaccard | Japanese | English | Jaccard |
| 1 | 覚える | learn | .219 | 単語 | word | .345 | 書く | write | .210 |
| 2 | 発音 | pronunciation | .117 | 覚える | learn | .263 | 見る | look | .103 |
| 3 | 聞く | listen | .096 | 見る | look | .141 | 意味 | meaning | .083 |
| 4 | 使う | use | .086 | 読む | read | .110 | 発音 | pronunciation | .079 |
| 5 | テスト | test | .082 | 音読 | reading aloud | .110 | 繰り返す | repeat | .074 |
| 6 | 質問 | question | .065 | 繰り返す | repeat | .102 | 綴り | spelling | .060 |
| 7 | TOEIC | TOEIC | .064 | テスト | test | .089 | 英語 | English | .050 |
| 8 | 英語 | English | .060 | ノート | notebook | .071 | 思い浮かべる | recall | .044 |
| 9 | 意識 | awareness | .048 | 使う | use | .071 | 紙 | paper | .044 |
| 10 | 考える | think | .047 | 勉強 | study | .054 | 授業 | class | .044 |

**Note. $N$ = 178; Higher values indicate stronger associations.**

Figure 4 shows a co-occurrence network illustrating the relationships between frequently occurring words and the strategy clusters. In the network, line thickness represents the strength of co-occurrence, and circle size indicates word frequency. Larger nodes located near the center of the network represent words commonly used across all clusters, suggesting that activities such as looking up words, writing, reading, and self-testing were widely reported by learners. The network also reveals lexical tendencies associated with each strategy cluster, providing a basis for the qualitative interpretations presented below.

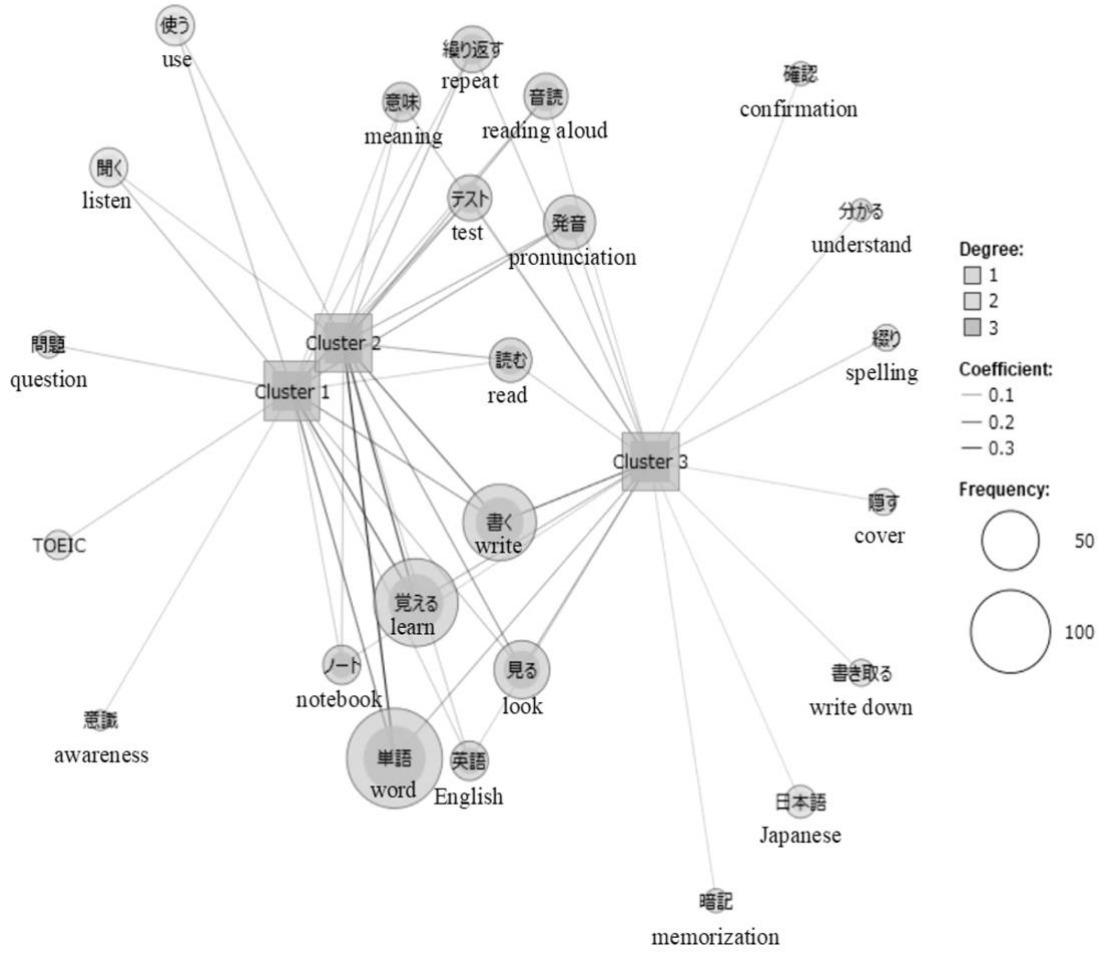


**Figure 4. Co-occurrence Network in Responses to Open-ended Question**

**Note. Node size represents word frequency. Edge thickness reflects co-occurrence coefficients. Japanese labels are followed by English translations.**

It should be noted that individual responses often contained multiple VLS (e.g., writing, repeating, pronouncing). In the examples below, italicized expressions highlight the lexical items most strongly associated with each cluster in the co-occurrence analysis. Cluster membership was determined based on overall co-occurrence patterns rather than the presence of a single word. Consequently, some responses may include strategies that also appear in other clusters.

Words frequently associated with Cluster 1 included *listen*, *use*, and *question*, suggesting more active and self-regulated engagement with VLS. Representative participant responses illustrating these tendencies are presented below (translated from the original Japanese, with cluster-related expressions italicized). These examples help contextualize the quantitative cluster distinctions.

- I memorize words and idioms by grouping them to some extent while looking for common points in words. I also apply the knowledge of affixes to the words that I learn, pronouncing and *listening* to them. (ID: S1002)
- In addition to learning by writing, I try to learn words that are difficult to remember by associating them with images. I also make sentences *using* the words that I learned. (ID: S1079)
- I learn words not only by writing in a notebook, but also by pronouncing words and asking *questions* with my friends. (ID: S1097)

Words frequently associated with Cluster 2 included *repeat, look*, and *reading aloud*. Representative excerpts from participants' responses illustrating these behaviors are provided below.

- I try to *repeat* reading and writing words until I can remember. (ID: S1025)
- I memorize the spelling and meaning of words by *looking* at them. I also learn pronunciation by *looking* at phonetic symbols. (ID: S1059)
- I *read aloud* the sentences in my word book together with the accompanying audio CD. (ID: S1083)

Words frequently associated with Cluster 3 included *write*, *spelling*, and *memorization*. Representative excerpts from participants in this group are presented below.

- I read the Japanese translations and *write* the words. (ID: S1040)
- I *write* the *spellings* of English words and their Japanese translations alternately. (ID: S1134)
- I *write* words down in my notebook until I remember them. I also use *memorization* apps. (ID: S1180)

Taken together, these examples illustrate how different clusters reflect distinct tendencies in learners' VLS, suggesting differences not only in strategy frequency but also in the depth and diversity of strategy engagement.

## Discussion

This study examined how VLS relate to learners' VS and vocabulary development over time among Japanese STEM-oriented upper secondary students at KOSEN. Using a mixed-methods design that integrated quantitative and qualitative approaches, the study explored these relationships from complementary perspectives. By combining Rasch-based VS estimation, cluster analysis, two-way mixed ANOVA, and text mining of open-ended responses, the study provides both theoretical and pedagogical insights into vocabulary development in Japanese EFL contexts.

Regarding RQ1, learners reported using a range of strategy types, with contextualizing and oral rehearsal emerging as the most frequently employed. These were followed, in descending order, by grouping, metacognitive, social, and traditional strategies. The prominence of contextualizing and grouping strategies may reflect engagement in deeper cognitive processes such as imagery, organization, encoding, and retrieval (Gu & Johnson, 1996; Horino & Ichikawa, 1997). In contrast, oral rehearsal and traditional strategies are more commonly associated with surface-level retention. Metacognitive strategies facilitate planning, monitoring, and evaluation of learning, while social strategies support collaborative vocabulary practice (Kojic-Sabo & Lightbown, 1999; Nyikos & Fan, 2007). Overall, these findings align with previous research indicating that effective vocabulary learners tend to employ a balanced combination of cognitive, metacognitive, and social strategies (Schmitt, 1997).

RQ2 explored patterns of VS growth in relation to VLS use. Learners were classified into high, medium, and low strategy-use groups, which differed primarily in the overall frequency and breadth of strategy engagement rather than in qualitatively distinct repertoires. High strategy users consistently reported applying a wider range of strategies, particularly metacognitive, contextualizing, and grouping strategies. Longitudinal analyses indicated that learners with higher levels of strategy engagement tended to show greater vocabulary growth over time, supporting earlier findings that strategic learning may contribute to vocabulary acquisition (Atay & Ozbulgan, 2007; Mizumoto & Takeuchi, 2009b). These results suggest the importance of sustained and reflective learning practices for vocabulary growth.

RQ3 examined whether learners with different levels of strategy engagement exhibited somewhat different longitudinal vocabulary growth trajectories. The observed interaction between strategy group and time suggests that learners with higher levels of strategy engagement tended to demonstrate greater cumulative vocabulary development. Although the magnitude of these differences was modest, incremental gains can become pedagogically meaningful when accumulated across extended instructional periods. This finding supports previous work linking frequent strategy use to vocabulary development (Fan, 2003; Gu, 2010; Schmitt, 1997) and extends Akase (2021) by providing longitudinal evidence that sustained strategy engagement is associated with gradual but enduring vocabulary growth. However, because VLS were measured at a single time point, cluster membership was determined post hoc, limiting causal interpretation and allowing for possible reverse causality. Future studies should therefore assess strategy use longitudinally to examine reciprocal relationships between VLS engagement and vocabulary development.

For RQ4, qualitative text-mining analyses revealed systematic differences in self-reported strategy use across groups, adding an interpretive layer to the quantitative findings. High strategy users frequently described phonological, organizational, and metacognitive practices, such as noticing affixes, imagining usage contexts, and engaging in peer interaction. Phonological practices specifically included attention to pronunciation and listening, suggesting engagement with phonological aspects of vocabulary learning beyond simple reading aloud or repetition. Medium strategy users relied mainly on word lists and reading aloud, with occasional references to active use strategies. In contrast, low strategy users focused primarily on rote memorization, translation, and orthographic matching. These patterns suggest that vocabulary development may be associated not only with strategy frequency but also with the diversity and depth of strategy use (Gu & Johnson, 1996; Sanaoui, 1995; Schmitt, 1997).

In summary, the findings suggest that learners with higher levels of strategy engagement and a broader repertoire of strategies tend to demonstrate greater vocabulary growth over time. The mixed-methods design enabled triangulation of quantitative and qualitative evidence,

strengthening the validity of the conclusions and clarifying the mechanisms underlying observed statistical relationships. From a pedagogical perspective, the results suggest that instruction should place greater emphasis on metacognitive awareness and deeper cognitive strategies, while also encouraging social interaction and contextualized vocabulary use. Beyond strategy frequency alone, future research may benefit from examining how individual differences—such as motivation and self-efficacy—interact with VLS engagement to influence long-term vocabulary development across diverse learning contexts (Akase, 2023).

## Conclusion

This longitudinal mixed-methods study explored how VLS relate to VS development among Japanese EFL learners in STEM-oriented KOSEN settings. By integrating three years of quantitative data with qualitative analyses of learners' self-reported strategies, the study provides a nuanced understanding of how differences in strategy engagement are associated with vocabulary growth over time.

The findings suggest that vocabulary development is related not only to the overall level of strategy use but also to qualitative differences in how learners engage with vocabulary. Learners showing higher strategic engagement tended to report more integrative and meaning-focused approaches, such as attending to word structure, imagining usage contexts, and collaborating with peers. In contrast, learners with lower engagement more often relied on mechanical practices, including memorizing word–translation pairs and orthographic repetition. Taken together, these patterns indicate that deeper and more diverse strategy repertoires are modestly but meaningfully associated with vocabulary outcomes over time.

However, several limitations warrant consideration. Vocabulary knowledge was assessed exclusively through receptive measures, and the reduced number of questionnaire items may have somewhat limited the precision of strategy measurement. In addition, although associations between strategy engagement and vocabulary development were observed, causal interpretations remain tentative, as learners with larger vocabularies may also be more inclined to adopt a wider range of strategies. Other learner variables, such as motivation, self-regulation, and self-efficacy, are also likely to contribute to vocabulary growth. Finally, the focus on Japanese STEM-oriented students and the imbalanced gender distribution limit the generalizability of the findings.

Despite these limitations, the study offers longitudinal mixed-methods evidence that sustained and diversified strategy engagement may be associated with vocabulary development in EFL classroom settings. Although individual effects appear modest, such differences may accumulate across instructional years, potentially leading to educationally meaningful outcomes. From a pedagogical perspective, the findings underscore the importance of fostering metacognitive awareness and encouraging deeper processing strategies, rather than relying primarily on rote memorization. Future research should examine how instructional interventions can support the development of flexible strategy repertoires and explore how strategy use interacts with motivational and affective factors across diverse learning contexts.

## Acknowledgements

The author wishes to express sincere gratitude to Dr. James Sick, whose insightful comments, suggestions, and encouragement remain invaluable to the development of this study. He thanks two anonymous reviewers for their constructive feedback and suggestions, which greatly improved the quality of this paper. He also acknowledges financial support from the National Institute of Technology, Nagano College.

## About the Author

**Masaki Akase** is an Associate Professor at the National Institute of Technology, Nagano College, Japan. His research focuses on vocabulary acquisition, vocabulary assessment and testing, vocabulary learning strategies, and individual differences in EFL learning. He is the author of *Longitudinal Growth in Vocabulary Size, Vocabulary Learning Strategy Use, Motivation, and Self-Efficacy in Japanese EFL Learners* (2023). ORCID: 0000-0001-7712-0578.

## To Cite this Article

Akase, M. (2026). The role of vocabulary learning strategies in longitudinal vocabulary growth. *Teaching English as a Second Language Electronic Journal (TESL-EJ), 30*(2). https://doi.org/10.55593/ej.30118a6

## References

Aizawa, K., & Mochizuki, M. (2010). *Eigo goi shido no jissen idea shu: Katsudourei kara test sakusei made [Practical handbook for English vocabulary teaching: From examples of activities to drawing on tests].* Taishukan Shoten.

Akase, M. (2021). A longitudinal analysis of growth in vocabulary size and strategy use: Focusing on Japanese EFL senior high school students. *Journal of Pan-Pacific Association of Applied Linguistics, 25*(2), 105–130. https://doi.org/10.25256/PAAL.25.2.6

Akase, M. (2022). Longitudinal measurement of growth in vocabulary size using Rasch-based test equating. *Language Testing in Asia, 12*(1), Article 5. https://doi.org/10.1186/s40468-022-00155-8

Akase, M. (2023). *Longitudinal growth in vocabulary size, vocabulary learning strategy use, motivation, and self-efficacy in Japanese EFL learners.* Independently published.

Anderson, R. C., & Freebody, P. (1981). Vocabulary knowledge. In J. T. Guthrie (Ed.), *Comprehension and teaching: Research reviews* (pp. 77–117). International Reading Association.

Atay, D., & Ozbulgan, C. (2007). Memory strategy instruction, contextual learning and ESP vocabulary recall. *English for Specific Purposes, 26*(1), 39–51. https://doi.org/10.1016/j.esp.2006.01.002

Beglar, D. (2010). A Rasch-based validation of the vocabulary size test. *Language Testing, 27*(1), 101–118. https://doi.org/10.1177/0265532209340194

Benesse Educational Research and Development Institute. (2019). *Kōichi-sei no eigo gakushū ni kansuru chōsa [Investigation of 1st year high school students' English learning].* https://benesse.jp/berd/up_images/research/part1.pdf

Boers, F. (2024). Bridging intentional and incidental vocabulary learning. In F. Teng. & B. L. Reynolds (Eds). Researching incidental vocabulary learning in a second language (pp. 182–195). Routledge. https://doi.org/10.4324/9781003270782

Bond, T., & Fox, C. (2001). *Applying the Rasch model: Fundamental measurement in the human sciences.* Lawrence Erlbaum Associates Publishers.

Bond, T., Yan, Z., & Heene, M. (2020). *Applying the Rasch Model: Fundamental measurement in the human sciences* (4th ed.). Routledge. https://doi.org/10.4324/9780429030499

Cameron, L. (2001). *Teaching languages to young learners.* Cambridge University Press. https://doi.org/10.1017/CBO9780511733109

Catalán, R. M. J. (2003). Sex differences in L2 vocabulary learning strategies. *International Journal of Applied Linguistics, 13*(1), 54–77. https://doi.org/10.1111/1473-4192.00037

Cohen, A. D. (2011). *Strategies in learning and using a second language*. Routledge.

Creswell, J. W. (2015). *A concise introduction to mixed methods research.* Sage.

Dörnyei, Z. (2005). *The psychology of the language learner: Individual differences in second language acquisition* (3rd ed.). Lawrence Erlbaum Associates.

Dörnyei, Z., & Dewaele, J.-M. (2022). *Questionnaires in second language research: Construction, administration, and processing* (3rd ed.). Routledge. https://doi.org/10.4324/9781003331926

Fan, M. (2003). Frequency of use, perceived usefulness, and actual usefulness of second language vocabulary strategies: A study of Hong Kong learners. *The Modern Language Journal, 87*(2), 222–241. https://doi.org/10.1111/1540-4781.00187

Griffiths, C. (2018). *The strategy factor in successful language learning: The Tornado Effect* (2nd ed.). Multilingual Matters. https://doi.org/10.21832/9781783099757

Gu, P. Y. (2010). Learning strategies for vocabulary development. *Reflections on English Language Teaching, 9*(2), 105–118.

Gu, P. Y. (2019). Strategies for learning vocabulary. In S. Webb (Ed.), The Routledge handbook of vocabulary studies (pp. 271–287). Routledge. https://doi.org/10.4324/9780429291586

Gu, P. Y., & Johnson, R. K. (1996). Vocabulary learning strategies and language learning outcomes. *Language Learning, 46*(4), 643–679. https://doi.org/10.1111/j.1467-1770.1996.tb01355.x

Hamada, A., Iso, T., Kojima, M., Aizawa, K., Hoshino, Y., Sato, K., Sato, R., Chujo, J., & Yamauchi, Y. (2021). Development of a vocabulary size test for Japanese EFL learners using the New JACET List of 8,000 Basic Words. *JACET Journal, 65*, 23–45. https://doi.org/10.32234/jacetjournal.65.0_23

Henning, G., Hudson, T., & Turner, J. (1985). Item response theory and the assumption of unidimensionality for language tests. *Language Testing, 2*(2), 141–154. https://doi.org/10.1177/026553228500200203

Higuchi, K. (2014). *Shakai chousa no tameno keiryou tekusuto bunseki [Quantitative text analysis for social researchers: A contribution to content analysis]*. Nakanishiya Shuppan.

Horino, M., & Ichikawa, S. (1997). Koukousei no eigo gakushu ni okeru gakushu doki to gakushu houryaku [Learning motives and strategies in high-school students' English learning]. *Japanese Journal of Educational Psychology, 45*(2), 140–147. https://doi.org/10.5926/jjep1953.45.2_140

JACET Basic Word Revision Committee (Ed.). (2003). *JACET list of 8000 basic words.* Japan Association of College English Teachers.

JASP Team. (2020). *JASP* (Version 0.12.2) [Computer software]. https://jasp-stats.org/

Kasahara, K. (2006). Producing a revised version of the Mochizuki Vocabulary Size Test. *JLTA Journal, 9*, 55–72. https://doi.org/10.20622/jltaj.9.0_55

Katagiri, K. (2009). A three-year longitudinal study of vocabulary size in Japanese SHS students and a description of their developmental patterns. *ARELE, 20*, 131–140. https://doi.org/10.20581/arele.20.0_131

Koizumi, R., & Mochizuki, M. (2011). Development and validation of the PC version of the Mochizuki Vocabulary Size Test. *JACET Journal, 53*, 35–55.

Kojic-Sabo, I., & Lightbown, P. M. (1999). Students' approaches to vocabulary learning and their relationship to success. *The Modern Language Journal, 83*(2), 176–192. https://doi.org/10.1111/0026-7902.00014

Lai, Y. (2016). EFL learners' vocabulary consolidation strategy use and corresponding performance on vocabulary tests. *Taiwan Journal of TESOL, 13*(1), 33–70. https://doi.org/10.30397/TJTESOL.201604_13(1).0002

Laufer, B., Elder, C., Hill, K., & Congdon, P. (2004). Size and strength: Do we need both to measure vocabulary knowledge? *Language Testing, 21*(2), 202–226. https://doi.org/10.1191/0265532204lt277oa

Laufer, B., & Goldstein, Z. (2004). Testing vocabulary knowledge: Size, strength, and computer adaptiveness. *Language Learning, 54*(3), 399–436. https://doi.org/10.1111/j.0023-8333.2004.00260.x

Lawson, M. J., & Hogben, D. (1996). The vocabulary-learning strategies of foreign-language students. *Language Learning, 46*(1), 101–135. https://doi.org/10.1111/j.1467-1770.1996.tb00642.x

Linacre, J. M. (2019). *Winsteps® Rasch measurement computer program* (Version 4.7) [Computer software]. Winsteps.com.

Meara, P. (1996). *The vocabulary knowledge framework*. Vocabulary Acquisition Research Group Virtual Library, Swansea University.

Meara, P. (2002). The rediscovery of vocabulary. *Second Language Research, 18*(4), 393–407. https://doi.org/10.1191/0267658302sr211xx

MEXT. (2018). *Kōtō gakkō gakushū shidō yōryō kaisetsu: Gaikokugo-hen [Explanatory comment for the new study of course guidelines for foreign language in senior high schools]*. Kairyūdō.

Milton, J., & Hopwood, J. (2023). *Vocabulary in the foreign language curriculum*. Routledge.

Mizumoto, A., & Takeuchi, O. (2009a). A closer look at the relationship between vocabulary learning strategies and TOEIC scores. *TOEIC Research Report, 4*, 1–34. The Institute for International Business Communication.

Mizumoto, A., & Takeuchi, O. (2009b). Examining the effectiveness of explicit instruction of vocabulary learning strategies with Japanese EFL university students. *Language Teaching Research, 13*(4), 425–449. https://doi.org/10.1177/1362168809341511

Mochizuki, M. (1998). *Nihonjin eigo gakushu-sha no tameno goi saizu tesuto* [A vocabulary size test for Japanese learners of English]. *The Institute for Research in Language Teaching Bulletin, 12*, 27–53.

Nakamura, T. (2002). *Vocabulary learning strategies: The case of Japanese learners of English*. Koyo Shobo.

Nation, I. S. P. (2006). *How large a vocabulary is needed for reading and listening? Canadian Modern Language Review, 63*(1), 59–82. https://doi.org/10.3138/cmlr.63.1.59

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

Nation, I. S. P. (2022). *Learning vocabulary in another language* (3rd ed.). Cambridge University Press. https://doi.org/10.1017/9781009093873

Nation, I. S. P., & Beglar, D. (2007). A vocabulary size test. *The Language Teacher, 31*(7), 9–13.

Nyikos, M., & Fan, M. (2007). A review of vocabulary learning strategies: Focus on language proficiency and learner voice. In A. D. Cohen & E. Macaro (Eds.), *Language learner strategies* (pp. 251–273). Oxford University Press.

Oxford, R. L. (1990). *Language learning strategies: What every teacher should know*. Newbury House.

Qian, D. D. (1999). Assessing the roles of depth and breadth of vocabulary knowledge in reading comprehension. *Canadian Modern Language Review, 56*(2), 282–307. https://doi.org/10.3138/cmlr.56.2.282

Rasch, G. (1960). *Probabilistic model for some intelligence and achievement tests.* Danish Institute for Educational Research.

Rasekh, Z. E., & Ranjbary, R. (2003). Metacognitive strategy training for vocabulary learning. *TESL-EJ, 7*(2), 1–17. https://tesl-ej.org/ej26/a5.html

Read, J. (2000). *Assessing vocabulary.* Cambridge University Press. https://doi.org/10.1017/CBO9780511732942

Sanaoui, R. (1995). Adult learners' approaches to learning vocabulary in second languages. *The Modern Language Journal, 79*(1), 15–28. https://doi.org/10.1111/j.1540-4781.1995.tb05410.x

Schmitt, N. (1997). Vocabulary learning strategies. In N. Schmitt & M. McCarthy (Eds.), *Vocabulary: Description, acquisition and pedagogy* (pp. 199–227). Cambridge University Press.

Schmitt, N. (2000). *Vocabulary in language teaching*. Cambridge University Press.

Schmitt, N. (2008). Review article: Instructed second language vocabulary learning. *Language Teaching Research*, *12*(3), 329–363. https://doi.org/10.1177/1362168808089921

Schmitt, N. (2010). *Researching vocabulary: A vocabulary research manual*. Palgrave Macmillan. https://doi.org/10.1057/9780230293977

Schmitt, N. (2014). Size and depth of vocabulary knowledge: What the research shows. *Language Learning, 64*(4), 913–951. http://dx.doi.org/10.1111/lang.12077

Schmitt, N., Nation, I. S. P., & Kremmel, B. (2019). Moving the field of vocabulary assessment forward: The need for more rigorous test development and validation. *Language Teaching, 53*(1), 109–120. https://doi.org/10.1017/S0261444819000326

Schmitt, N., & Schmitt, D. (2014). A reassessment of frequency and vocabulary size in L2 vocabulary teaching. *Language Teaching, 47*(4), 484–503. https://doi.org/10.1017/S0261444812000018

Stæhr, L. S. (2009). Vocabulary knowledge and advanced listening comprehension in English as a foreign language. *Studies in Second Language Acquisition, 31*(4), 577–607. https://psycnet.apa.org/doi/10.1017/S0272263109990039

Stewart, J. (2014). Do multiple-choice options inflate estimates of vocabulary size on the VST? *Language Assessment Quarterly, 11*(3), 271–282. https://doi.org/10.1080/15434303.2014.922977

Tabachnick, B. G., & Fidell, L. S. (2013). *Using multivariate statistics* (6th ed.). Pearson.

Teng, F., & Uchihara, T. (2024). Factors affecting incidental vocabulary learning. In F. Teng & B. L. Reynolds (Eds.), *Researching incidental vocabulary learning in a second language* (pp. 131–148). Routledge. https://doi.org/10.4324/9781003270782-12

Tseng, W.-T., & Schmitt, N. (2008). Toward a model of motivated vocabulary learning: A structural equation modeling approach. *Language Learning, 58*(2), 357–400. https://doi.org/10.1111/j.1467-9922.2008.00442.x

Webb, S., & Nation, P. (2017). *How vocabulary is learned*. Oxford University Press.

Webb, S., & Paribakht, T. S. (2015). What is the relationship between the lexical profile of test items and performance on a standardized English proficiency test? *English for Specific Purposes, 38*(1), 34–43. https://doi.org/10.1016/j.esp.2014.11.001

Webb, S., Sasao, Y., & Ballance, O. (2017). The updated vocabulary levels test: Developing and validating two new forms of the VLT. *ITL – International Journal of Applied Linguistics, 168*(1), 33–69. https://doi.org/10.1075/itl.168.1.02web

Wilkins, D.A. (1972). *Linguistics in language teaching.* Arnold.

Zhang, X., & Lu, X. (2015). The relationship between vocabulary learning strategies and breadth and depth of vocabulary knowledge. *The Modern Language Journal, 99*(4), 740–753. https://doi.org/10.1111/modl.12277

## Appendix A

### Sample Items from Vocabulary Size Test 4 (Rasch-based)

Select English word which represents the meaning of each Japanese one from (1) to (6) and write the number on the answer sheet (Originally in Japanese).

1. (1,000-Word Level)

| | | | | | |
|---|---|---|---|---|---|
| 1. 丸い入れ物 | | | 2. クッションのある長いす | | |
| (1) bath | (2) lamp | (3) phone | (4) pot | (5) sofa | (6) stove |
| 2. 歓迎する，喜んで迎える | | | 3. 働く | | |
| (1) help | (2) receive | (3) show | (4) touch | (5) welcome | (6) work |
| 5.だいだい色の果物 | | | 6.トウモロコシ | | |
| (1) chicken | (2) corn | (3) orange | (4) plant | (5) tomato | (6) vegetable |
| 7. 食事 | | | 8. １つ，１個，１片 | | |
| (1) air | (2) meal | (3) piece | (4) sign | (5) sound | (6) while |
| 9. 気がつく，見つける | | | 10. ～になる | | |
| (1) become | (2) clean | (3) explain | (4) notice | (5) set | (6) wear |
| 11.引く，引っ張る | | | 12.見る | | |
| (1) lend | (2) look | (3) pull | (4) read | (5) tell | (6) worry |
| 13. 不可解なこと，不思議なこと | | | 14. 試験 | | |
| (1) act | (2) butterfly | (3) exam | (4) mystery | (5) tennis | (6) trouble |
| 15. 船 | | | 16. 手 | | |
| (1) deer | (2) goal | (3) ship | (4) butter | (5) hand | (6) park |
| 17. おじ | | | 18. 手袋 | | |
| (1) mother | (2) season | (3) uncle | (4) bird | (5) giant | (6) glove |
| 19. 質問 | | | 20. 記録 | | |
| (1) movie | (2) record | (3) red | (4) interview | (5) machine | (6) question |
| 21. すてきな，すばらしい | | | 22. 大きい | | |
| (1) cool | (2) hot | (3) large | (4) least | (5) nice | (6) quiet |
| 23. 落とす | | | 24. 手に持つ | | |
| (1) wink | (2) follow | (3) drop | (4) believe | (5) hold | (6) put |
| 25. 彼女のもの | | | 26. 私の | | |
| (1) below | (2) hers | (3) my | (4) past | (5) which | (6) whom |

2. (3,000-Word Level)

| | | | | | |
|---|---|---|---|---|---|
| 1. 運動 | | | 2. 肉，肉体 | | |
| (1) beach | (2) campaign | (3) economy | (4) flesh | (5) glory | (6) worker |
| 3. 知らない人 | | | 4. 知事，統治者 | | |
| (1) bulk | (2) companion | (3) governor | (4) prayer | (5) relief | (6) stranger |
| 5. わずかな部分，分数 | | | 6. 軍隊，集団 | | |
| (1) fare | (2) fraction | (3) intention | (4) jaw | (5) solution | (6) troop |
| 7. 儀式 | | | 8. 緊急事態 | | |
| (1) apparatus | (2) boundary | (3) ceremony | (4) emergency | (5) horizon | (6) sympathy |
| 9. 雇うこと，雇用 | | | 10. 事務員 | | |
| (1) clerk | (2) compound | (3) confidence | (4) employment | (5) invention | (6) oxygen |

| | | | | | |
|---|---|---|---|---|---|
| 11. 世帯 | | | 12. 襟 | | |
| (1) collar | (2) household | (3) impression | (4) inquiry | (5) opposition | (6) remark |
| 13. 港 | | | 14. 奴隷 | | |
| (1) advantage | (2) arrow | (3) guidance | (4) harbor | (5) reform | (6) slave |
| 15. 改定する | | | 16. 腐る，朽ちる | | |
| (1) decay | (2) distribute | (3) fasten | (4) fold | (5) isolate | (6) revise |
| 17. 生き物，動物 | | | 18. 事情，状況 | | |
| (1) circumstanc | (2) creature | (3) liberty | (4) presence | (5) sequence | (6) unit |
| 19. 気がついて | | | 20. まっすぐに立っている | | |
| (1) absent | (2) aware | (3) central | (4) drunk | (5) historical | (6) upright |
| 21. 動機 | | | 22. 梯子 | | |
| (1) county | (2) fame | (3) ladder | (4) motive | (5) sphere | (6) tongue |
| 23. 調査，測量 | | | 24. 賢いこと，知恵 | | |
| (1) construction | (2) deed | (3) exception | (4) fortune | (5) survey | (6) wisdom |
| 25. 実際に | | | 26. とにかく | | |
| (1) actually | (2) anyhow | (3) completely | (4) indeed | (5) somewhere | (6) whenever |

3. (5,000-Word Level)

| | | | | | |
|---|---|---|---|---|---|
| 1. 洗面器，鉢 | | | 2. 縫い目 | | |
| (1) basin | (2) buckle | (3) flake | (4) hive | (5) seam | (6) spear |
| 3. 魔女 | | | 4. 子孫，子ども | | |
| (1) awe | (2) guardian | (3) offspring | (4) pioneer | (5) shepherd | (6) witch |
| 5. 日用品 | | | 6. 打撲傷，打ち身 | | |
| (1) bruise | (2) commodity | (3) interference | (4) propaganda | (5) strategy | (6) token |
| 7. 優れていること | | | 8. 群れ | | |
| (1) excellence | (2) furnace | (3) kidney | (4) nap | (5) swarm | (6) thermometer |
| 9. そらす | | | 10. (力を) 出す，行使する | | |
| (1) conserve | (2) divert | (3) exert | (4) heave | (5) recede | (6) shudder |
| 11. 衣服 | | | 12. 装飾品，飾り | | |
| (1) ditch | (2) garment | (3) ornament | (4) peninsula | (5) profile | (6) usage |
| 13. いたずら，わるさ | | | 14. 回復すること，復旧 | | |
| (1) claw | (2) collision | (3) courtesy | (4) epoch | (5) mischief | (6) restoration |
| 15. ぐるぐる巻く，渦巻く | | | 16. 追い払う，撃退する | | |
| (1) blink | (2) chuckle | (3) heighten | (4) repel | (5) sob | (6) whirl |
| 17. 目的，意図 | | | 18. 国籍 | | |
| (1) deputy | (2) familiarity | (3) intent | (4) nationality | (5) transplant | (6) validity |
| 19. 花が咲く，花をつける | | | 20. 叱る | | |
| (1) bloom | (2) compute | (3) erase | (4) mislead | (5) scold | (6) slam |
| 21. 遺伝子の，遺伝的な | | | 22. 半狂乱となった，血迷った | | |
| (1) adjective | (2) collective | (3) considerate | (4) eloquent | (5) frantic | (6) genetic |
| 23. 診断する | | | 24. 蒸発する | | |
| (1) certify | (2) diagnose | (3) evaporate | (4) expel | (5) inject | (6) merge |
| 25. 統計的な，統計上の | | | 26. もうけになる，有利な | | |
| (1) agreeable | (2) clinical | (3) honorable | (4) profitable | (5) prospective | (6) statistical |

4. (2,000-Word Level)

| | | | | | |
|---|---|---|---|---|---|
| 1. 報酬，ほうび | | | 2. 敵 | | |
| (1) comparison | (2) decision | (3) enemy | (4) feather | (5) influence | (6) reward |
| 3. 約束する | | | 4. 論じ合う，議論する | | |
| (1) check | (2) collect | (3) discuss | (4) hurry | (5) miss | (6) promise |
| 5. 洪水 | | | 6. 設備，備品 | | |
| (1) account | (2) courage | (3) equipment | (4) factor | (5) flood | (6) lack |
| 7. しつけ，鍛錬 | | | 8. 海岸 | | |
| (1) benefit | (2) coast | (3) discipline | (4) division | (5) soap | (6) truth |
| 9. 必要な | | | 10. 眠って | | |
| (1) asleep | (2) fierce | (3) lazy | (4) necessary | (5) personal | (6) violent |
| 11. 著者，作者 | | | 12. 血，血液 | | |
| (1) author | (2) blood | (3) choice | (4) growth | (5) item | (6) luck |
| 13. 直す，繕う | | | 14. 含む | | |
| (1) contain | (2) defend | (3) delay | (4) mend | (5) occur | (6) trace |
| 15. 予期する，期待する | | | 16. 勝つ，勝利する | | |
| (1) achieve | (2) expect | (3) possess | (4) remove | (5) suggest | (6) win |
| 17. 不平を言う，苦情を言う | | | 18. 指し示す，示す | | |
| (1) complain | (2) depend | (3) imagine | (4) indicate | (5) protect | (6) split |
| 19. 汚い，よごれた | | | 20. 軍事的な，軍事の | | |
| (1) angry | (2) calm | (3) dirty | (4) final | (5) military | (6) secret |
| 21. 好奇心の強い | | | 22. 生の，加工していない | | |
| (1) curious | (2) equal | (3) independent | (4) raw | (5) social | (6) steady |
| 23. 分かれた，分離した | | | 24. 緊急の，差し迫った | | |
| (1) bright | (2) frequent | (3) initial | (4) safe | (5) separate | (6) urgent |
| 25. (損害などを) こうむる，受ける | | | 26. 口論する，論ずる | | |
| (1) argue | (2) blame | (3) describe | (4) filter | (5) join | (6) suffer |

5. (4,000-Word Level)

| | | | | | |
|---|---|---|---|---|---|
| 1. 電球，眼球 | | | 2. 軌道 | | |
| (1) bulb | (2) cluster | (3) cricket | (4) lounge | (5) mode | (6) orbit |
| 3. 認めること，承認 | | | 4. 祝宴，宴会 | | |
| (1) admission | (2) bull | (3) feast | (4) geometry | (5) hedge | (6) succession |
| 5. 調査，研究，捜査 | | | 6. 展覧会，展示会 | | |
| (1) constitution | (2) contribution | (3) definition | (4) exhibition | (5) investigation | (6) resistance |
| 7. 強烈さ，激しさ | | | 8. 地理 | | |
| (1) geography | (2) intensity | (3) mortgage | (4) payment | (5) privilege | (6) removal |
| 9. 船舶 | | | 10. 部分，一部 | | |
| (1) cereal | (2) craft | (3) deposit | (4) pastry | (5) portion | (6) registration |
| 11. 割り当てる | | | 12. つぶやく，ぶつぶつ言う | | |
| (1) assign | (2) cope | (3) grasp | (4) mutter | (5) reckon | (6) retreat |
| 13. 気品，威厳 | | | 14. 賞，賞品 | | |
| (1) award | (2) cottage | (3) dignity | (4) microscope | (5) tenant | (6) tragedy |

| | | | | | |
|---|---|---|---|---|---|
| 15. 大勝利，大成功 | | | 16. 名声，威信 | | |
| (1) contempt | (2) distinction | (3) passion | (4) prestige | (5) realm | (6) triumph |
| 17. 仲直りさせる | | | 18. 邪魔する，妨げる | | |
| (1) conclude | (2) hinder | (3) murmur | (4) reconcile | (5) stagger | (6) weave |
| 19. 購入する，買う | | | 20. 再び始める | | |
| (1) alternate | (2) collapse | (3) fetch | (4) pat | (5) purchase | (6) resume |
| 21. 火山 | | | 22. 知人，知り合い | | |
| (1) acquaintance | (2) enthusiasm | (3) landscape | (4) victim | (5) volcano | (6) window |
| 23. 単数の | | | 24. ことばの，言語の | | |
| (1) administrative | (2) atomic | (3) concrete | (4) frank | (5) linguistic | (6) singular |
| 25. 大胆な | | | 26. 堅い，厳格な | | |
| (1) bold | (2) dumb | (3) false | (4) mature | (5) rigid | (6) suspicious |

6. (6,000-Word Level)

| | | | | | |
|---|---|---|---|---|---|
| 1. ローマ教皇 | | | 2. とさか | | |
| (1) butt | (2) crest | (3) dormitory | (4) pope | (5) ration | (6) sulfur |
| 3. 合併，合同 | | | 4. 下水道 | | |
| (1) bidder | (2) merger | (3) predator | (4) premium | (5) sewer | (6) wedge |
| 5. 用心して，慎重に | | | 6. 代わる代わる，交替に | | |
| (1) acutely | (2) alternately | (3) cautiously | (4) periodically | (5) simultaneously | (6) traditionally |
| 7. 頻繁に，絶えず | | | 8. 精神的に，知能的に | | |
| (1) continually | (2) correctly | (3) forcibly | (4) historically | (5) legally | (6) mentally |
| 9. 糖尿病 | | | 10. 妊娠中絶 | | |
| (1) abortion | (2) artillery | (3) coffin | (4) diabetes | (5) diagnosis | (6) immunity |
| 11. かたわらへよける | | | 12. 深くする，濃くする | | |
| (1) deepen | (2) dissatisfy | (3) imprint | (4) pinpoint | (5) shuffle | (6) sidestep |
| 13. 包む，覆う | | | 14. (水などを) どっと流す | | |
| (1) dodge | (2) envelop | (3) flop | (4) flush | (5) perch | (6) sip |
| 15. (尾を) 振る | | | 16. 偽装する，作り上げる | | |
| (1) churn | (2) engulf | (3) forge | (4) poise | (5) smear | (6) wag |
| 17. 混乱させる，分裂させる | | | 18. 起訴する，告発させる | | |
| (1) brood | (2) clog | (3) disrupt | (4) distrust | (5) prosecute | (6) wade |
| 19. 沼地，湿地 | | | 20. 立法府，立法機関 | | |
| (1) coral | (2) legislature | (3) librarian | (4) pitfall | (5) potter | (6) swamp |
| 21. 違法の，嫡出でない | | | 22. 早い，時期尚早の | | |
| (1) attentive | (2) feeble | (3) illegitimate | (4) lame | (5) premature | (6) resistant |
| 23. 潮の | | | 24. 弱い，衰弱した | | |
| (1) feeble | (2) frontal | (3) literal | (4) surgical | (5) tidal | (6) tribal |
| 25. 経済的に，節約して | | | 26. 根本的に，本質的に | | |
| (1) aptly | (2) economically | (3) fundamentally | (4) genuinely | (5) immensely | (6) radically |

## Appendix B

### Sample Items from the Vocabulary Learning Strategies Questionnaire

What strategy do you use to learn or memorize English words and phrases? The following sentences describe some of the learning strategies. Please circle the number that you think is most applicable (Originally in Japanese).

| | | |
|---|---|---|
| 1. Strongly disagree | 2. Disagree | 3. Partly disagree. |
| 4. Partly agree | 5. Agree | 6. Strongly agree |

1. I memorize a word by connecting it to my personal experience (e.g., I'm shy but cheerful, etc.).
2. I consciously set aside time to study vocabulary in order to prepare for tests such as TOEIC, TOEFL or STEP EIKEN: English Proficiency Test.
3. I study vocabulary with the intention of writing and speaking English.
4. I make flashcards to practice words (e.g., write English words on one side and the Japanese meanings on the back).
5. I study the meaning and usage of words with a friend.
6. I imagine whether a word has a good (positive) or a bad (negative) meaning to try to remember it.
7. I remember new words by checking where they appear in the study materials.
8. I remember words with similar spellings or related meanings together.
9. I remember words by saying them out loud repeatedly.
10. I try to expose myself to English vocabulary by reading or listening to English a lot.
11. I group words together to memorize them (e.g., fruit = apple, orange, strawberry, pear, etc.).
12. I make my own study plan before learning new vocabulary (e.g., "I'm going to learn 10 words a day.").
13. I study out loud each example sentence in which the word I want to learn is used.
14. I write the word many times to remember it.
15. I talk to a foreign teacher (native speaker) to learn the meaning and usage of a word.
16. I ask my friend to test me on words to see if I have learned them.
17. I remember a word by associating it with similar expressions or paraphrases (e.g., shout = to talk very loudly).
18. I study the words out loud, paying attention to how they are pronounced.
19. I memorize a word with an image of an action or object that can be associated with the meaning of the word (e.g., when learning "seize", I use my hands to grip something.).
20. I keep a vocabulary book or word list (e.g., a commercial vocabulary book) to check the vocabulary anytime I wish.
21. I consciously study the part of speech of new words (e.g., nouns, verbs, adjectives, adverb, etc.) to remember them.
22. I set aside time for vocabulary study by testing myself on words to see if I have learned them.
23. I associate new words with the ones I already know (e.g., when learning "monkey", I relate it to other words such as "banana", "sweet", "yellow", "jungle", etc.).
24. I try to manage the learning environment so as to expose myself to English vocabulary.
25. I use colors and highlighters to mark new important words in a text.
26. I ask a teacher for the spelling and meaning of words to remember them.
27. I make a mental image of the spelling of a word I want to remember.
28. I make vocabulary notebooks, add information to them, and use them as a reference.
29. I review the words I have learned regularly outside of class or school.
30. I associate the word I know with the synonyms (e.g., begin and start) or antonyms (e.g., good and bad) to remember them.
31. I use rhyme to remember a word (e.g., take = mot<u>teiku</u>, tot<u>teiku</u>, tshure<u>teiku</u>).
32. I write down new words in my notebook immediately.
33. I study vocabulary through media such as movies, TV, radio, and the Internet.