

# Vocabulary Education & Research Bulletin

Winter 2018

Volume 7, Issue 2

We bring you the new issue of the VERB, still aglow from an event-filled autumn with a great SIG forum at JALT national, and before that a successful and inspiring Vocabulary Symposium at Meiji Gakuin, where it seemed that the voices of the vocabulary greats were still reverberating from the walls of Vocab@Tokyo, held there not so long ago. This issue has two poster presentation synopses from the symposium. **Joshua Brook Antle** brings us a replication study of Jiang's (2009) look at consciousness-raising tasks for collocation acquisition. Then **Jean-Pierre Joseph Richard** has built a bank of written productive NGSL items and provided initial validation. But first, in a short paper, **Andrew Blaker** does an analysis of several major *tango-cho* vocabulary notebooks used by high school students in Japan.

And this issue also welcomes the VERB's new co-editor, Tim Stoeckel, who has long been active with both the VERB and the Vocabulary SIG. We have gotten off to a good start and are looking forward to another vocabulary research-filled year in 2019, especially with Vocab@Leuven coming up in July. More information can be found in the SIG News section. Best wishes to all!

VERB editors

Magda Kitano & Tim Stoeckel

## Table of Contents

### Short Papers

Andrew Blaker \_\_\_\_\_ 2

*Vocabulary Coverage of High School Vocabulary Notebooks*

### Poster Presentation Synopses

Joshua Brook Antle \_\_\_\_\_ 9

*Noticing Collocations in Language: Investigating the Effectiveness of  
Consciousness-Raising Tasks for Acquiring Collocational Knowledge*

Jean-Pierre Joseph Richard \_\_\_\_\_ 14

*Developing Written Productive Items for an Institution-Level NGSL Assessment  
Program*

SIG News \_\_\_\_\_ 21

**Short Paper**

## Vocabulary Coverage of High School Vocabulary Notebooks

Andrew Blaker    andrew@hassaku.nz

The way in which English is taught and studied at high schools in Japan has an impact on the language ability of university entrants, who many readers of this journal teach. In the previous issue of *VERB*, Kitano (2018) noted that McLean, Hogg and Kramer (2014) found an unexpected gap in knowledge of high frequency words in students with both high and low vocabulary size scores. This paper investigates the notebooks that are used by high school students to study new vocabulary and the words they contain in order to ascertain a possible origin of this gap. A vocabulary coverage analysis of these notebooks is conducted using the British National Corpus (BNC) High Frequency Word List (HFWL) (The British National Corpus, 2007), and the New General Service List (NGSL) (Browne, Culligan, & Phillips, 2013). Finally, unique vocabulary items between the books are identified, and the results discussed.

### Notebooks

Vocabulary notebooks (also known as *tango-chō*) are vocabulary study books commonly used by junior and senior high schools in Japan as a tool for English vocabulary acquisition. They consist of lists of English words with Japanese translations. Often, but not always, an example sentence is given in both English and Japanese. The Ministry of Education (MEXT) guidelines for English language education state that students be introduced to 1,200 words at junior high school level, and a further 1,800 words during high school (Mizumoto, 2010). MEXT does not, however, provide a list of which words are to be taught at any level (Kaneko, 2014). Vocabulary notebooks are not the sole source of vocabulary input for the students. However, as the notebooks provide a simple single source of learning for the students, they are also a source of testing material for the teachers. As a result, teachers use these notebooks to guide students in self-study, and weekly word tests are created using the assigned words.

Furthermore, the National Center (university entrance) Test heavily influences the words chosen by the books, with each of the notebooks indicating either by section or by annotation which words most commonly appear in the test. The notebooks are often available in 1,800, 3,000, and 4,500-word varieties. Guidelines currently recommend that

a high school graduate should understand 3,000 words (MEXT, 2009), however MEXT does not specify whether this number is based on word families or lemmas. Based on this recommendation, schools often opt for the 3,000-word variety. Therefore, this study investigates four 3,000-word notebooks selected based on availability and popularity of usage in the author's experience in high schools.

In addition to the four vocabulary notebooks, the Oxford 3000 vocabulary list is also used. This is a list compiled by the Oxford Learner's Dictionary as a "list of the 3000 most important words to learn in English" ("The Oxford 3000," 2018, para. 1), and serves as a comparison of material serving the same purpose.

Table 1 introduces each notebook and shows the number of vocabulary items and total number of words used in each of the notebooks examined.

Table 1

*Vocabulary Items in Each Notebook and in the Oxford 3000 List*

Notebook Name	Vocabulary Items	Total Words
<i>Database 3000</i>	1600	1900
<i>Corpus 3000</i>	1473	2570
<i>System 3000</i>	2000	2881
<i>Phrase で英単語 3000</i>	1533	2245
<i>Oxford 3000</i>	3000	3000

*Note.* Total Words includes idioms, synonyms, and opposites.

As Table 1 shows, these notebooks contain between 1,473 and 2,000 vocabulary items, significantly fewer than 3,000 words.

## Method

The analysis is conducted using three different methods: coverage of the BNC HFWL (henceforth the HFWL), distribution of the NGSL, and enumeration of unique words. The HFWL is a list of the 14,000 most frequently identified word families from the BNC, divided into 1,000-word levels by Nation (2006). The NGSL is a list of 2,801 modified lemmas designed as useful high frequency vocabulary words for students of English as a second language divided into 1<sup>st</sup>, 2<sup>nd</sup>, and 3<sup>rd</sup> 1,000-word levels. The Oxford 3000 is a similar list to the NGSL designed by Oxford and uses a mix of both lexeme and word family (Burkett, 2015). This study regards word family as the counting unit for ease of comparison to these commonly available lists.

Comprehension suffers without fluency in high-frequency words (Nation, 2013). Different vocabulary coverage studies propose differing levels of coverage required for adequate comprehension, with 95% (Nation, 2001) and 98% (Hu & Nation, 2000; Schmitt, Jiang, & Grabe, 2011) commonly accepted values. Therefore for a vocabulary list to provide enough words for adequate comprehension, when compared to a piece of writing, 95% or 98% of words in the reading would be found in the vocabulary list. This is described as 95% or 98% coverage. To be easily comparable to other studies, this study includes both coverage rates of 95% and 98% in its analysis. The vocabulary lists were extracted from the notebooks, and the analysis consists of comparing these lists to each of the 1,000-word levels of the HFWL using AntWordProfiler (Anthony, 2014) to establish coverage rates of 95%, 98%, and 100%. The notebooks' vocabulary lists are then compared to the NGSL lists, to indicate what proportion of the words come from each NGSL level. Lastly, the notebooks' lists are compared to each other to identify the number of unique items between the books.

**Results**

*HFWL Coverage*

The first analysis shows the basic coverage of the vocabulary lists generated from the notebooks. This shows how far through each of the 1000-word levels of the HFWL are required to cover 95%, 98%, and 100% of each notebook. This is shown in Figure 1.

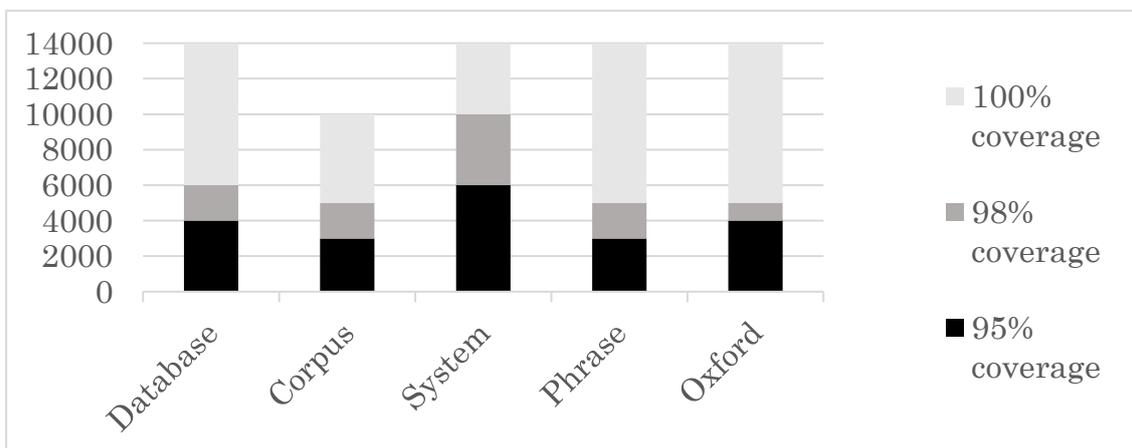


Figure 1. HFWL Coverage

Figure 1 shows that learners would need knowledge of the first 3,000 or 4,000 word families of the HFWL to reach 95% coverage of the target words presented in all but one of the notebooks. Given that the recommended vocabulary level of a high school

graduate is 3,000 words, this is level appropriate. If the coverage is lifted to 98%, however, knowledge of 5,000 or 6,000 HFWL entries is required. The biggest difference lies in the remaining 2%, in which most of the books contain words up to the 14,000-word level. *Corpus*, however, contains words only as high as the 10,000-word level. As it claims to be based on the BNC HFWL, this is most likely deliberate planning on the publisher's part.

### *NGSL Distribution*

The NGSL is now commonly used as a method of determining vocabulary levels in Japanese universities (Stoeckel & Bennett, 2015), and the next coverage analysis is conducted using the three bands of the NGSL.

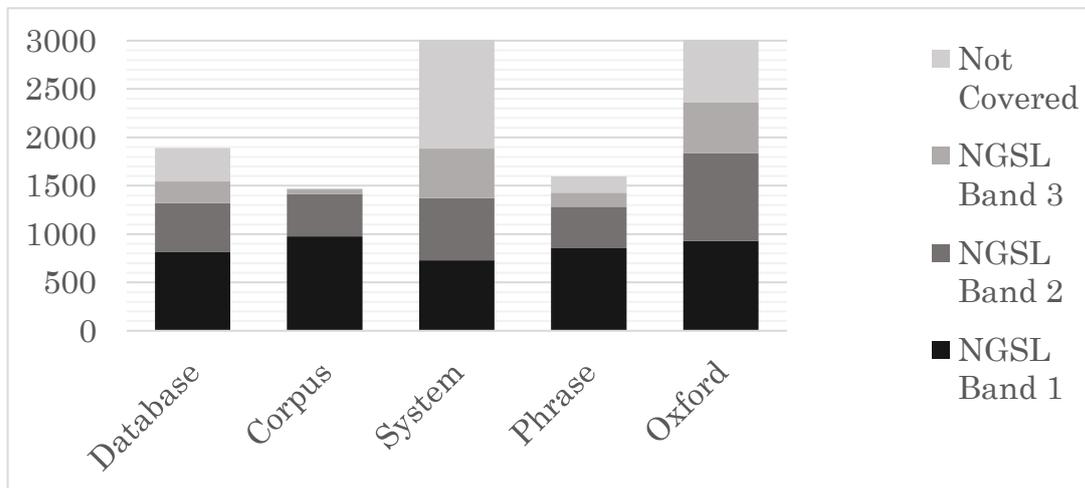


Figure 2. Number of Entries from Vocabulary Notebooks and the Oxford 3000 List that Are in the NGSL

As would be expected, Figure 2 shows that the 1<sup>st</sup> 1,000 is most represented, the 2<sup>nd</sup> 1,000 less so, and the 3<sup>rd</sup> band (2,001 through 2,801) the least represented. The *System* book and the Oxford list are notable for the large number of words from the 3<sup>rd</sup> band and the number of words that do not appear in the NGSL, with this largely being due to the differing sizes of the lists.

### **Differences between Words Presented**

The final analysis conducted looks at the difference in the words presented in each of the lists generated from the notebooks. Table 2 shows how many words contained in the two compared books are unique; in other words it is an indication of how many words are not shared. It should be noted that these comparisons are conducted using the raw, un-

lemmatized vocabulary lists, and as such the numbers are expected to be larger than if lemmatized lists were used.

Table 2

*Differences in Words Presented*

	<i>Corpus</i>	<i>System</i>	<i>Phrase</i>
<i>Database</i>	840	1386	784
<i>Corpus</i>		1010	664
<i>System</i>			1144

The most significant comparison apparent in Table 2 is between *Database* and *System*. While only containing 1,900 and 2,881 words respectively, 1386 unique words were found. The smallest number of unique words found was between *Corpus* (2570 words) and *Phrase* (2245 words), and even given these smaller vocabulary list sizes, 664 unique items were identified.

**Discussion**

The analysis conducted using the BNC HWFL showed that to achieve a 95% coverage rate of the vocabulary notebooks used in this study requires the top 3,000 most commonly used words in English, which suits the expectation of MEXT for high school students. Ninety eight percent requires 5,000 or more words, which is considerably larger than the high school students' expected levels. The most surprising result was the remaining 2% coverage for most of the notebooks required up to the 14,000-word level. While in a native reading environment this is nothing unusual, these books are designed to teach high school students English. It raises the question of why these low-frequency words are being introduced at a high school level.

The large number of unique words found between books is the greatest cause for concern. This shows that the vocabulary students are introduced to is significantly shaped by the notebook chosen by the school. It may not be necessary for each student's learning experience to be identical, but if MEXT were to provide a basic list of required words, preferably based on a high-frequency list, it would go a long way toward providing a more uniform and useful language learning experience for students.

This paper has investigated one possible source of the word gap identified by Kitano (2018) and identified two issues with notebooks introducing vocabulary at the high school level: vocabulary item choice based on high-frequency usage, and

inconsistency of vocabulary between books. These issues and their effects bear further investigation.

## References

- Anthony, L. (2014). *AntWordProfiler* (Version 1.4.1) [Computer Software]. Tokyo: Waseda University. Available from: [www.laurenceanthony.net/software](http://www.laurenceanthony.net/software).
- Browne, C., Culligan, B., & Phillips, J. (2013). The New General Service List. Retrieved from <http://www.newgeneralservicelist.org>.
- Burkett, T. (2015). An investigation into the use of frequency vocabulary lists in university intensive English programs. *International Journal of Bilingual & Multilingual Teachers of English*, 3(2).
- Hu, M. and Nation, I. S. P. (2000) Unknown vocabulary density and reading comprehension. *Reading in a Foreign Language*, 13(1), 403-430.
- Kaneko, M. (2014). Is the vocabulary level of the reading section of the TOEFL internet-based test beyond the lexical level of Japanese senior high school students? *Vocabulary Learning and Instruction*, 3(1), 44-50.
- Kitano, M. (2018). The high frequency word gap: What words don't students know? *Vocabulary Education & Research Bulletin*, 7(1), 2-5.
- MEXT (2009). *高等学校学習指導要領 [The Course of Study Guidelines for Upper Secondary School]*. Retrieved from: [http://www.mext.go.jp/a\\_menu/shotou/new-cs/youryou/kou/kou.pdf](http://www.mext.go.jp/a_menu/shotou/new-cs/youryou/kou/kou.pdf).
- McLean, S., Hogg, N., & Kramer, B. (2014). Estimations of Japanese university learners' English vocabulary sizes using the vocabulary size test. *Vocabulary Learning and Instruction*, 3(2), 47-55.
- Mizumoto, A. (2010). *Exploring the art of vocabulary learning strategies: A closer look at Japanese EFL university students*. (Doctoral Dissertation). Retrieved from <http://hdl.handle.net/10112/9084>
- Nation, I. S. P. (2001). *Learning vocabulary in another language*. Cambridge: Cambridge University Press.
- Nation, I. S. P. (2006). How large a vocabulary is needed for reading and listening? *Canadian Modern Language Review*, 63(1), 59-82.
- Nation, I. S. P. (2013). Vocabulary acquisition in second language acquisition. In C. A. Chapelle (Ed.), *The Encyclopedia of Applied Linguistics*. Oxford: Wiley-Blackwell.
- Schmitt, N., Jiang, X., & Grabe, W. (2011). The percentage of words known in a text and reading comprehension. *The Modern Language Journal*, 95(1), 26-43.

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

The British National Corpus (2007). Distributed by Bodleian Libraries, University of Oxford, on behalf of the BNC Consortium. Retrieved from <http://www.natcorp.ox.ac.uk/>

The Oxford 3000 (2018). *Oxford Learner's Dictionaries*. Retrieved from <https://www.oxfordlearnersdictionaries.com/about/oxford3000>

**Poster Presentation Synopsis****Noticing Collocations in Language: Investigating the Effectiveness of Consciousness-Raising Tasks for Acquiring Collocational Knowledge**

Joshua Brook Antle    joshua.antle@gmail.com

This study was a replication of an earlier investigation conducted by Jiang (2009) in which she used consciousness-raising tasks to improve productive use of English collocations. Two goals of consciousness-raising tasks are to aid in the retention of target vocabulary items and to empower learners to notice aspects of the L2 in future input. Lewis (2000) states, “the conscious noticing of features of the language that learners meet does facilitate acquisition” (p. 158).

**Background and Aims**

For the improvement of productive ability, collocations are an especially important aspect of word knowledge. A collocational focus will lead to improvements in both spoken and written fluency (Hill, 2000; Lewis, 2000). Researchers believe that collocations help speakers by reducing the processing load thereby making it easier for them to create a comprehensible utterance (Schmitt & Carter, 2004; Wood, 2010; Wray, 2000; Wray & Perkins, 2000).

While research on collocations has gained momentum, many language learners remain unaware of the importance these multi-word structures have on their productive ability and how to properly acquire these structures. Collocations are problematic for language learners for several reasons including L1 interference (Fan, 2009; Hashemi & Eskandari, 2017; Ozaki, 2011) and overuse of a few general items (Ying & O’Neill, 2009). While there are difficulties in teaching these structures, the potential benefits are worth effort (Bahns & Eldaw, 1993; Hill, 2000; Jiang, 2009). A consciousness-raising approach may be an efficient way to both help the language learners acquire target collocations within a class and prepare these students to notice collocations in reading and listening texts in the future.

Jiang’s study (2009) revealed that for Chinese university-aged English language learners it is necessary to raise awareness of collocations and that teaching materials should address collocations while incorporating the findings of corpus research. My study investigates whether these findings will be consistent in a Japanese context. The

following research question was investigated:

After a semester of consciousness-raising tasks, will the learners better understand the nature of collocations and their importance for language learning?

### **Sample Population**

This study was conducted at a public university in Japan. In total, 32 students agreed to participate. Due to small class sizes, the data were collected over the course of three years; each year, a one-semester class was used for the treatment. The students were all at a similar level, low-intermediate, with TOEFL scores ranging from 380 to 470.

### **Methods**

The procedure used was similar to Jiang's study (2009). Towards the end of the weekly class, the students were given a reading text of approximately 250 words and were encouraged to read it as quickly as possible. The students recorded their starting and finishing time, so they could calculate their reading speed. They then answered several comprehension questions about the text. For homework, they completed a series of follow-up exercises which included:

- writing down good expressions from the text they want to learn,
- creating original sentences using collocations from the text,
- finding other common collocates for nouns from the text, and
- completing sentence-level cloze exercises using target collocations from the text but in different contexts.

At the beginning of the following class, the students compared their homework answers and then retold the story using the expressions they had written down as a guide. This pattern was repeated for ten classes, and this treatment was presented to the students as part of their normal course work. After the treatment phase, the students completed a questionnaire which had ten Likert scale statements and six open-ended questions. For the purposes of this paper, the findings will be limited to the Likert scale results. For each of the questionnaire's ten statements, the participants indicated their level of agreement: strongly agree (SA), agree (A), not sure (NS), disagree (D) and strongly disagree (DD). The ten statements are as follows:

1. Memorizing word clusters helps me towards target-like.

2. There is not much difference between memorizing single words and word clusters.
3. Now I jot down a word and the company it keeps because I have recognized it is very important to remember word clusters.
4. Collocations are often very easy words, but often I am not sure which words can go with which words.
5. Simply remembering many individual words does not really mean one is good in English.
6. If I was not asked to jot down the good expressions, I may have overlooked some collocations.
7. My first language, Japanese, influences my collocation use in English to some extent.
8. It helps to constantly remind myself of avoiding possible Japanese-like collocations when I use English.
9. I am sure I will be more competent in collocation if I keep on jotting down good expressions.
10. The inclusion of collocation tasks in teaching materials is a very effective way to help me to become a better language user.

## **Results**

To allow for comparison, the questionnaire items are the same as in Jiang's (2009) study. Prior to administering the questionnaire, the teacher explained any potentially difficult phrases within the statements. The results for these statements are shown in Table 1. To allow the calculation of a mean ( $M$ ) and standard deviation ( $SD$ ), each level of agreement was assigned a numerical value: strongly agree (5), agree (4), not sure (3), disagree (2) and strongly disagree (1).

All of the students who initially agreed to take part in this study completed the questionnaire. There were sporadic absences over the course of the ten treatment lessons; however, as attendance was a part of their final grade, the majority of the participants attended all of the lessons.

The findings indicated that the consciousness-raising approach was successful in the learners' opinion. Specifically, the participants indicated that acquiring greater collocational knowledge will help them towards more native-like productive ability ( $M = 4.28$ ). Furthermore, the relatively high average score for items seven ( $M = 3.87$ ) and eight ( $M = 3.84$ ) suggested the students became aware of the negative influence their L1 can have on collocational knowledge. The students also revealed in statement three ( $M = 4.16$ )

how they were more likely to focus on collocations in the input they are exposed to instead of simply focusing on the individual words.

Collocational knowledge is an important aspect of vocabulary depth. The responses for items four ( $M = 3.88$ ) and five ( $M = 3.94$ ) showed the learners were aware of this facet of vocabulary knowledge. It was also noteworthy the high average score for item ten ( $M = 4.25$ ) which focused on the value of including collocation tasks in teaching materials.

Table 1

*Participants' Views on Collocation and Consciousness-raising Tasks*

Q#	N	SA	A	NS	D	DD	M	SD
1	32	10	21	1			4.28	0.52
2	32	6	10	9	7		3.47	1.05
3	32	9	19	4			4.16	0.63
4	32	8	15	7	1	1	3.88	0.94
5	32	9	15	5	3		3.94	0.91
6	31	4	12	13	2		3.58	0.81
7	31	2	24	4	1		3.87	0.56
8	31	5	17	8	1		3.84	0.73
9	32	3	22	5	2		3.81	0.69
10	32	15	10	7			4.25	0.80

### Conclusions and Future Directions

To summarize, consciousness-raising tasks are an effective way to improve collocational knowledge for university-level language classes. Additionally, in the participants' opinion, these tasks increase the likelihood of them noticing collocations in the future and are a worthwhile addition to language teaching materials. These findings are consistent with those of Jiang's (2009) study despite the different teaching context. Possible limitations for this study include the comparatively small number of participants and that the data were collected in three separate semesters over the course of three years. Jiang's (2009) study had 75 participants and collected all data within one 12-week semester.

Future research should investigate whether the noticing skills acquired by the learners are still actively being used. The students indicated that they believe this vocabulary learning strategy is beneficial, but since a lack of collocational competence

hinders learners' productive ability more than their receptive ability, they will have to make a conscious effort to notice collocations in the input to which they are exposed.

### References

- Bahns, J., & Eldaw, M. (1993). Should we teach EFL students collocations? *System*, 21(1), 101–114. doi:10.1016/0346-251X(93)90010-E
- Fan, M. (2009). An exploratory study of collocational use by ESL students: A task based approach. *System*, 37, 110–123.
- Hashemi, M. R., & Eskandari, R. (2017). The learning of congruent and incongruent collocations utilizing dynamic assessment. *The Language Teacher*, 41(6), 9-14.
- Hill, J. (2000). Revising priorities: From grammatical failure to collocational success. In M. Lewis (Ed.), *Teaching Collocation: Further Developments in the Lexical Approach* (pp. 47–69). Hove, England: Language Teaching Publications.
- Jiang, J. (2009). Designing pedagogic materials to improve awareness and productive use of L2 collocations. In A. Barfield and H. Gyllstad (Eds.), *Researching Collocations in Another Language* (pp. 99–113). New York: Palgrave Macmillan.
- Lewis, M. (2000). Learning in the lexical approach. In M. Lewis (Ed.), *Teaching Collocation: Further Developments in the Lexical Approach* (pp. 155 – 185). Hove, England: Language Teaching Publications.
- Ozaki, S. (2011). Teaching collocations effectively with the aid of L1. *The Language Teacher*, 35(3), 37.
- Schmitt, N., & Carter, R. (2004). Formulaic sequences in action. In N. Schmitt (Ed.), *Formulaic Sequences* (pp. 1–22). Amsterdam: John Benjamins.
- Wood, D. (2010). *Formulaic language and second language speech fluency*. London: Continuum.
- Wray, A. (2000). Formulaic sequences in second language teaching: Principle and practice. *Applied Linguistics*, 21, 463–489.
- Wray, A., & Perkins, M. R. (2000). The functions of formulaic language: an integrated model. *Language and Communication*, 20(1), 1–28.
- Ying, Y., & O'Neill, M. (2009). Collocation learning through an “AWARE” approach: Learner perspectives and learning process. In A. Barfield and H. Gyllstad (Eds.), *Researching Collocations in Another Language* (pp. 181–193). New York: Palgrave Macmillan.

**Poster Presentation Synopsis****Developing Written Productive Items for an Institution-Level NGSL Assessment Program**

Jean-Pierre Joseph Richard    richard.jean-pierre@u-nagano.ac.jp

**Background**

Students at the University of Nagano, a new public university which opened in April 2018, must demonstrate receptive and productive mastery of the New General Service List (NGSL; Browne, Culligan, & Phillips, 2013) before a required second-year study abroad program. As one criterion to demonstrate this, quarterly written receptive and productive vocabulary tests are being developed. Receptive items were selected from the New General Service List Test (NGSLT; Stoeckel & Bennett, 2015). Productive items similar to those on the Productive Vocabulary Levels Test (PVL; Laufer & Nation, 1999) were generated for this project. This paper summarizes the development of the productive items (Study 1) and reports on an initial test that combined the receptive and productive items (Study 2).

The NGSLT is a test of receptive vocabulary knowledge. It contains 100 items, 20 per 560-word level of the NGSL (Stoeckel & Bennett, 2015). NGSLT items employ a multiple-choice format with specifications similar to those of the Vocabulary Size Test (VST; Nation & Beglar, 2007).

The PVL was developed to measure written controlled productive vocabulary knowledge (Laufer & Nation, 1999). *Controlled* implies that there are constraints on word use. The PVL cues a one-word fill-in-the-blank response in a sentence (e.g., The nu\_\_\_\_ was helping the doctor in the operating room). The correct response, “nurse,” is cued by “nu,” which prevents test-takers from responding with other semantically appropriate responses. However, there are several known problems with the PVL, including the fact that background knowledge or knowledge of word parts can influence test results (Mochizuki & Aizawa, 2001). Nonetheless, due to institutional requirements that receptive and productive knowledge of the NGSL be tested, PVL-type items were developed.

**Aims**

The aims of the present study were twofold. The first was to build a bank of written productive NGSL items. The second was to provide initial validation evidence for a test

which combined these productive items with receptive items from the NGSLT. Construct validity was assessed on the basis of Rasch analysis and on the correlation between scores on the productive and receptive sections of the test. The minimum acceptable Pearson correlation was set at .4 (see Richard, 2018).

### **Study 1: Methods and Results**

Productive items were written for 8-14 randomly selected words from each 100-word band of the NGSL. Where possible, high frequency vocabulary was used when writing the items. Specifically, 96.6% of the 2101 tokens for these items were from the first 2000 words of the NGSL. See the Appendix for the remaining words. In addition, most responses required base, rather than inflected, forms. Finally, prompts were worded as statements rather than questions. In total, 253 items were written and tested. They assessed 110 nouns (43%), 107 verbs (42%), and 36 adjectives, adverbs, and prepositions (14%).

Data were gathered at two convenience-sample sites whose ranks (*hensachi*) were near the expected rank of the University of Nagano. Sakura University (*hensachi* = 55) is a private, female-only university in Tokyo, where participants ( $n = 25$ ) were enrolled in a university-wide first-year English course. Kanto Seaside University (*hensachi* = 54) is a large, comprehensive private university located near Tokyo, where participants ( $n = 42$ ) were third- and fourth-year English majors enrolled in year-long academic writing courses. Multiple test forms were used at both sites to maximize the number of tested items, with a total of 146 items (58% of items) linking sites. Across sites, an average of 24.8 participants ( $SD = 12.1$ , Max = 61, Min = 7) completed each item. Following Elwood (2011), items were scored 2, 1, or 0 points for a correct, partially correct, or incorrect answer, respectively. Correct spelling was required for a score of 2. For partially correct answers, egregious errors were tolerated, such as *inturodakshon* for *introduction*. See Table 1 for a summary of methods and results.

Rasch analysis was conducted using WINSTEPS 3.81.0 software (Linacre, 2007). Infit or outfit mean square (MNSQ) values greater than 1.5 was the criterion for item underfit (Linacre, 2006). Five items were deleted from the model due to perfect scoring (*made*,  $n = 9$ ; *during*,  $n = 41$ ; *issue*,  $n = 32$ ; *involves*,  $n = 29$ ; *belong*,  $n = 12$ ). For the remaining 248 items, infit and outfit MNSQ values ranged from .59 to 1.52 and .31 to 3.34, respectively. After three misfitting persons were removed from the data set, 15 items were found to underfit the Rasch model (Table 2).

Table 1

*Summary of Methods and Results for Study 1*

	Sakura ( <i>n</i> = 25)	Kanto Seaside ( <i>n</i> = 42)
Weekly study target	160 headwords	250 headwords
Data collection (weeks)	Apr-Dec 2016 (18)	Sep 2016-Jan 2017 (11)
Items / week	6-10	10-15
Forms / week	3	3
Common items	3-6	6-8
Approximate time / test	5 minutes	10 minutes
<i>M</i> ( <i>SD</i> ) test-takers / item	15.69 (6.64)	15.17 (7.02)
Max (Min) test-takers / item	25 (7)	39 (7)
Raw score item difficulty <i>M</i> ( <i>SD</i> )	.62 (.25)	.55 (.26)
Compensation	5% of term grades	5% of term grade

*Note.* Common items refers to items that were on the forms at both sites.

Table 2. *Item Fit Statistics with MNSQ values above 1.50 after Deleting Misfitting Persons*

Item	<i>n</i>	Infit MNSQ	Outfit MNSQ
Pick	30	1.27	3.43
Value	16	1.18	3.08
Combination	12	1.18	2.38
Equipment	13	1.39	2.05
Force	7	1.53	1.97
Develop	24	1.16	1.94
Proud	12	1.27	1.83
Race	9	1.14	1.81
Behave	10	1.36	1.72
Measured	29	1.35	1.68
Crime	22	.98	1.63
Just	54	1.44	1.63
Contrast	11	1.74	1.52
Distinguish	21	1.19	1.62
Altogether	19	1.42	1.52

## Study 2: Methods and Results

In December 2016, participants ( $n = 98$ ) at Hongo College (*hensachi* = 52), a two-year college in Chūbu, completed a test that included 40 receptive items selected from the NGSALT and 45 productive items selected from those described in Study 1. Based on data from Study 1, the selected productive items covered a range of difficulty (CHIPs<sup>1</sup>  $M = 50.00$ ,  $SD = 0.92$ ,  $Max = 51.57$ ,  $Min = 46.80$ ). Study 2 participants were not given word lists to study for the test.

The receptive items had infit and outfit MNSQ values, ranging from .82 to 1.14 and .30 to 1.43, respectively. The productive items had initial infit and outfit MNSQ values ranging from .70 to 1.32 and .31 to 2.02, respectively. Seven items (*people*, 1.52; *like*, 1.64; *appeal*, 1.66; *argue*, 1.69; *post*, 1.88; *public*, 2.01; *country*, 2.02) had outfit MNSQ values above 1.50. After five misfitting persons were removed, infit and outfit MNSQ values ranged from .70 to 1.20 and .29 to 1.82, respectively. Only *post* remained underfitting to the Rasch model (outfit MNSQ = 1.82). This item was considered a candidate for deletion; however, after deleting it and rerunning Rasch analysis most parameter estimates did not improve. Thus, this item was retained.

The correlation between the receptive and productive Rasch person measures was large ( $r = .64$ ). The range of the receptive and productive items was satisfactory with a span of 5.31 and 5.87 CHIPs, respectively. Mean Rasch person ability measures for the receptive and productive items were 51.41 and 49.91, respectively, indicating that the former were somewhat easy and the latter were nearly evenly matched for this group. Considering the *hensachi* of Hongo College, the average person ability measures found with the current participants indicate that the test might be appropriate for post-secondary students at average-ranked Japanese institutions, although more data collection and analyses are needed. The analysis also indicated that both sets had high item reliability (receptive = .95, productive = .97), although person reliability was lower (receptive = .71, productive = .84). Importantly, both sets were found to be unidimensional as indicated by the low unexplained variance for the receptive (2.6 eigenvalues) and productive (2.9) items in the first contrast of a principal components analysis of Rasch residuals (Linacre, 2007). See Tables 3 and 4 for selected Rasch statistics and the principal components analysis of Rasch residuals.

---

<sup>1</sup> In the current study, person ability and item difficulty are delineated in a user-friendly, response probability scaling unit, known as CHIPs (Wright & Stone, 1979, p. 201). One CHIP is equal to .22 logits, with the mean of the scale set at 50.

Table 3

*Selected Rasch Statistics for Study 2 (NGSL 1-2801)*

	Receptive ( $k = 40$ )	Productive ( $k = 45$ )
Rasch item reliability	.95	.97
Rasch person reliability	.71	.84
Item separation (strata statistic)	4.23 (5.97)	6.05 (8.40)
Person separation (strata statistic)	1.57 (2.43)	2.30 (3.40)
Item difficulty span (range) in CHIPs	5.31 (47.21-52.52)	5.87 (46.29-52.16)
Person ability span (range) in CHIPs	4.74 (49.09-53.83)	3.90 (47.85-51.75)
Mean person ability in CHIPs	51.41	49.91

Table 4

*Principal Components Analysis of Rasch Residuals for the Receptive and Productive NGSL Test Sections*

	Receptive ( $k = 40$ )	Productive ( $k = 45$ )
Total raw variance in observations	60.0	84.1
Raw variance explained by measures	20.0	39.1
Raw variance explained by persons	6.6	9.1
Raw variance explained by items	13.4	29.9
Raw unexplained variance (total)	40.0	45.0
Unexplained variance in 1 <sup>st</sup> contrast	2.6	2.9

*Note.* All values are expressed in eigenvalues.

**Preliminary Conclusions and Future Directions**

This paper's aims were twofold: to develop a bank of items similar to those on the PVLТ and to use those items together with existing NGSLT items to develop a vocabulary test that assesses both written receptive and productive NGSL mastery for curricular purposes at the University of Nagano. Study 1 described the development of the productive items. Study 2 described the initial validation evidence for a vocabulary test that combined these productive items with NGSLT items. Preliminary results indicated that the different sets of items were unidimensional, they were reliable, and there was a large correlation between the scores on the receptive and productive sections of the test. Taken together, these results provide tentative support for construct validity and for the

continuance of this project. Although not detailed here, data from two quarterly vocabulary tests have already been collected at the University of Nagano, covering the first half of the NGSLS, with plans to collect data on the second half in the near future.

### References

- Browne, C., Culligan, B., & Phillips, J. (2013). *The New General Service List*. Retrieved from <http://www.newgeneralservicelist.org>.
- Elwood, J. A. (2011). *Enriching structural models of L2 willingness to communicate: The role of personality, ego permeability, and perceived distance*. Unpublished doctoral dissertation, Temple University, Philadelphia.
- Laufer, B., & Nation, P. (1999). A vocabulary-size test of controlled productive ability. *Language Testing*, 16, 33-51. doi:10.1177/026553229901600103
- Linacre, M. (2006). *A user's guide to WINSTEPS*. Chicago, IL: Winsteps.com
- Linacre, J. M. (2007). WINSTEPS (Version 3.81.0) [Rasch-model computer program] . Chicago, IL: Winsteps.com
- Mochizuki, M., & Aizawa, K. (2001). A validity study of the vocabulary size test of controlled productive ability. *Reitaku University Journal*, 73, 85-102.
- Nation, I. S. P., & Beglar, D. (2007). A vocabulary size test. *The Language Teacher*, 31(7), 9-13.
- Richard, J.-P. J. (2018). *Academic capital, learner goals, achievement orientations, perceptions of English, and effort: A mixed-methods study*. Unpublished doctoral dissertation, Temple University, Philadelphia.
- Stoeckel, T., & Bennett, P. (2015). A test of the New General Service List. *Vocabulary Learning and Instruction*, 4(1), 1-8. doi: 10.7820/vli.v04.1.stoeckel.bennett
- Wright, B. D. & Stone, M. H. (1979). *Best test design: Rasch measurement*. Chicago, IL: MESA Press.

**Appendix****Tokens in Study 1 Beyond the 2000 Most Frequent NGSL Headwords**NGSL3: Headword (tokens)

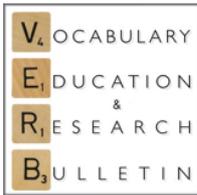
bread (1)	festival (1)	overseas (1)
cake (1)	fourth (1)	pen (2)
chocolate (1)	Friday (1)	penalty (1)
cream (1)	friendship (1)	three (1)
demonstration (1)	jacket (1)	two (4)
entrance (1)	magic (1)	
essay (2)	mom (1)	

NAWL: Headword (tokens)

anti (1)	homework (1)	regime (1)
artwork (1)	lab (1)	vocabulary (1)
elementary (1)	questionnaires (1)	

OFFLIST: Headword (tokens)

actress (1)	internet (2)	sore (1)
beef (1)	itchy (1)	spaghetti (1)
bicycle (1)	lottery (1)	tablet (1)
birthday (1)	mild (1)	tsunami (1)
blog (1)	oh (1)	typhoon (1)
cellphone (1)	ouch (1)	unaware (1)
disagreement (1)	peacefully (1)	unfair (1)
earthquake (2)	quit (1)	vanilla (1)
fantasy (1)	repay (1)	wallet (1)
flour (1)	robbery (1)	wisely (1)
fog (1)	rude (1)	yen (1)
gang (1)	sleepy (1)	
homeless (1)	smartphone (1)	



## **SIG News**

### **Upcoming Events**

#### **The JALT PanSIG Conference**

Konan University, Nishinomiya, Hyogo: May 18-19, 2019

The Vocabulary SIG is proudly sponsoring the following two events:

1. The Vocabulary SIG Forum entitled *New Corpus Tools and Methods for Data-Driven Learning in the Technical Writing Classroom*. This hands-on workshop will be conducted by Professor Laurence Anthony of Waseda University.
2. A presentation entitled *A Brief Update on NGSL Wordlist Research* by Professor Charles Browne of Meiji Gakuen University.

#### **The vocab@Leuven International Conference**

Leuven, Belgium: July 1-3, 2019

The Vocab@Leuven is the third Vocab@ conference. It aims to bring together researchers from different disciplines who investigate the learning, processing, teaching, and testing of second/foreign language vocabulary.

The call for abstracts is now closed. For more information:

<https://vocabatleuven.wordpress.com/>

#### **The 2019 Vocabulary SIG Symposium**

Waseda University, Tokyo: October 12, 2019

The 8<sup>th</sup> annual SIG symposium is a full day of presentations and discussion, including a poster session, followed by an evening of good food and conversation. Tune into the SIG website or Facebook feed in the coming months for more information.

**VERB Reviewers:** Phil Bennett, Thuy Bui, Tomoko Ishii, Brandon Kramer, Mimi Masson, Atsushi Mizumoto, John Racine, Rachel Ruegg, Jeff Stewart, Raymond Stubbe, and Yuka Yamamoto.

## VERB Call for Papers

---

The VERB welcomes submissions related to vocabulary research and education.

**Short papers** are peer reviewed, and may require rewriting and resubmission for acceptance. They must not exceed 1500 words, excluding references and titles. Short papers fall into the categories of completed research, ongoing research, and teaching and learning in practice.

**Other submissions** encouraged are classroom activities related to vocabulary, book reviews, opinion pieces, and event reports and commentary. All submissions are expected to adhere to APA 6th edition formatting guidelines.

Deadline for next issue: **March 15, 2019.**

For submissions and all correspondence: <[jaltvocabsig.verb@gmail.com](mailto:jaltvocabsig.verb@gmail.com)>

Latest information: <https://jaltvocab.weebly.com/publications.html>

The following are guidelines for short paper submissions (please include these sections):

---

<b>Completed research:</b>	<b>Ongoing research:</b>	<b>Teaching and learning in practice:</b>
* Background	* Background	* Theoretical framework
* Aims	* Aims	* Sample population
* Methods	* Methods	* Procedure
* Sample	* Sample	* (Preliminary) Results
* Results	* (Preliminary) Results	* (Preliminary) Conclusions
* Conclusions	* (Preliminary) Conclusions	* Future directions
* Future directions	* Future directions	

---

\*\*If you are thinking about submitting, however your article doesn't fit into one of the above categories, please email us at the above address and let us know what you would like to submit and we can work it out.

## ***Vocabulary Learning & Instruction* Call for Papers**

---

The Vocabulary SIG's *Vocabulary Learning and Instruction* (VLI) journal is calling for submissions for an upcoming issue. Submissions will be published online upon acceptance, and combined into an issue later this year.

VLI accepts long-form research papers (2000-7000 words) and brief reports, summaries, and commentaries (2000-3000 words) related to vocabulary acquisition, pedagogy, assessment, and lexical networks.

As an open journal, content is indexed on Google Scholar and made freely available on the internet without paywalls. Authors are free to also make their work available on sites such as academia.edu and researchgate.

All submissions are subject to a 2-step peer-review process:

A) Editors review manuscripts to ensure basic requirements are met, and that the work is of sufficient quality to merit external review. This process typically takes 1-2 weeks, at which point authors are informed of the outcome.

B) Submissions which meet these requirements are sent out for blind peer review by 2-3 experts in the field. This process takes approximately 1-2 months. Following external review, authors are sent copies of external reviewers' comments and notified of decisions (accept, accept pending changes, revise and resubmit, or reject).

Please see <http://vli-journal.org/submissions.html> for details