An Edited Version of the First Eight 1,000-Word Frequency Bands of the Japanese-English Version of the Vocabulary Size Test

Stuart McLean
Osaka Jogakuin

Tomoko Ishii
Meiji Gakuin University

Tim Stoeckel
University of Niigata Prefecture

Phil Bennett

Yuko Matsumoto
Miyazaki International College

This paper provides and explains the criteria by which the first eight 1,000-word frequency bands of the Japanese bilingual Vocabulary Size Test (VST) were revised. The VST (Nation & Beglar, 2007) was designed as a measure of vocabulary size for language learners. It was originally produced and validated in a monolingual format, but in recent years several bilingual versions have also been made. These variants may yield more accurate results, because they avoid conflating vocabulary knowledge with ability to decode answer choices in the L2. However, they have received little scrutiny beyond initial piloting and may therefore benefit from further examination and refinement (Nguyen & Nation, 2011). This paper describes the revision of the first eight 1,000-word frequency bands of the Japanese bilingual VST with the goal of increasing the test's unidimensionality and accuracy. The revisions (a) removed English loanwords from the answer choices to prevent examinees from correctly responding through phonological matching alone, (b) ensured that the parts of speech of each answer choice were identical, and (c) matched the lengths of answer choices.

本論文は、日本語版2ヶ国語 Vocabulary Size Test (VST) における最初の8000語レベルまでの改訂された単語出現頻度帯の基準を説明したものである。VST (Nation and Beglar, 2007) は言語学習者の語彙サイズを計測するために作られたもので、当初は英語のみで制作されていたが、近年様々な言語対応の版も作られている。これらの版では、英語の読解力に左右されることなくテストを受けられ、より正確な結果が得られることが予想される。しかしながら、それらの版は初期のパイロットテストの範囲を超える精査が殆どされていないため、更なる検証を重ねる事には意義がある (Nguyen & Nation, 2011)。本論文は、日本語版 VST の最初の8000語レベルの単語出現頻度帯について、測定の一次元性と正確さの向上を目的に改訂を試み、その結果をまとめたものである。改訂作業においては、 (a) 音による推測が容易なカタカナ語を選択肢から排除し、(b) 各選択肢の品詞を統一し、(c) 各選択肢の長さの均衡をとった。

The Monolingual VST

The Vocabulary Size Test (VST) (Nation & Beglar, 2007) was developed to provide estimates of the number of words that foreign or second language learners can recognize in a written format. Various monolingual forms of the VST exist, all of which are based on a frequency count of word families derived from the British National Corpus (BNC; Nation, 2006). The first twelve 1,000-word lists of the BNC were revised using word family range and frequency figures from the 10-million token spoken section of the BNC, the remaining 1,000 word lists bands are based on the entire 100-million token BNC. The most well-known version of the VST is the original 140-item version, developed to provide a comprehensive measure of vocabulary size from the first to the fourteenth 1,000-word families of English. More recently, two parallel 100-item forms covering the first 20,000-word families have also become available. For learners of low or intermediate proficiency, abbreviated forms have been used. In his validation study of the test, Beglar (2010) used an 80-item version measuring knowledge of the first eight 1,000-word bands with intermediate level students and a 40-item version measuring knowledge of the first four 1,000-word bands with low proficiency learners. This was because using lower frequency bands with these learners would have increased the chances of overestimating vocabulary size due to guessing. Examples of words from lower frequency bands are *fens, cranny, aperitif, soliloquy, talons* and *erythrocyte* at the ninth, tenth, eleventh, twelfth, thirteen and fourteenth 1,000-word bands, respectfully.

Elgort (2013) also took the position that it is not always optimal to estimate vocabulary size using the entire instrument. McLean, Hogg, and Kramer (2014) and McLean, Kramer, and Stewart (2015) provided insight into appropriate test length in the Japanese context. These researchers collected VST scores from Japanese university students and, based on subsequent retrospective interviews and statistical analysis, posited that the vocabulary sizes of most of these learners were overestimated when using the results from the first eight 1,000-word frequency bands. It may be concluded that for accurate estimates of vocabulary size from all but the most able of Japanese university students, utilization of bands beyond the first 8,000 words is best avoided. For this reason, we have limited the scope of our revisions of the Japanese VST to the first eight sections of the test.

Bilingual Versions of the VST

A number of bilingual versions of the VST have been developed, including Russian, Korean, Vietnamese, Mandarin, Persian, and Japanese. In these variants, each question stem is presented in English while the answer choices are presented in the learners' L1. Here is an example from the Japanese bilingual VST:

see: They **saw** it.
  a. 切った (cut)
  b. 待った (waited for)
  c. 見た (saw)
  d. 始めた (started)

Studies of the Vietnamese, Persian, and Russian versions of the VST have found that each provides reliable results, generally separating participants according to overall proficiency, and that the difficulty of items forms a continuum based on frequency (Elgort, 2013; Karami, 2012; Nguyen & Nation, 2011).

Relative Benefits and Limitations of Monolingual and Bilingual Versions of the VST

Unidimensionality

A fundamental purpose of test design is to create instruments which yield unidimensional results. In other words, test makers hope to isolate and measure a single target construct to the greatest degree possible. Bilingual tests play a key role in this. Elgort (2013), Karami (2012) and Nguyen and Nation (2011) argue that bilingual tests more accurately estimate vocabulary sizes of examinees with limited knowl-

edge of grammar and syntax. The monolingual versions of the VST include relative clauses (e.g., 'argued against the facts that supported it') and adjective/noun/verb phrases (e.g., 'full of energy'). Karami (2012) states:

> Complex grammatical knowledge as well as good reading skills are required on the part of the test takers to get the item right. It poses a problem for monolingual vocabulary size tests as these tests cannot be claimed to be tests of pure vocabulary knowledge. Other factors, such as grammatical knowledge and reading ability, contaminate the measurement (p. 55).

Bilingual versions of the VST may also reduce the level of anxiety examinees experience while completing the test (Elgort, 2013). Further, through the perception of relative ease, examinees' self-efficacy specific to the task of completing a bilingual VST may increase. Our level of self-efficacy influences "whether certain (coping) behaviors will be initiated, how much effort will be expended, and how long it will be sustained in the face of obstacles and aversive experiences" (Bandura, 1977, p. 191).

Time Necessary to Complete the Test

Owing to students' faster L1 reading speed relative to their L2, and the fact that bilingual tests allow for shorter answer choices, the bilingual VST can be completed more quickly than the monolingual form. Karami (2012) states that "long distractors are unavoidable in monolingual tests to completely convey the meaning of more difficult words" (p. 56). Shorter answer choices in the bilingual format may reduce the degree to which fatigue influences VST scores, particularly among less able readers, as such learners often have difficulty concentrating for more than 20 minutes when conducting L2 reading (Day & Bamford, 1998).

The Utilization of Only High Frequency Words in Answer Options

In the monolingual VST, answer choices were written with a restricted set of vocabulary. The answer choices of items testing the first two 1,000-word frequency bands were written using only the first 1,000 words of West's (1953) General Service List (GSL), while items from the third 1,000-word band onward were written using only the GSL. This raises two issues. First, examinees with very limited lexical knowledge might not be able to demonstrate understanding of items testing the first 1,000 words if they lack knowledge of words used in the answer choices. Second, accurately representing the

meaning of target words with very limited vocabulary can be problematic. For example, *hikidashi,* the direct Japanese translation of *drawer,* represents a much more accurate representation of *drawer* than "sliding box," the definition offered on a version of the monolingual test.

Revising the Bilingual Japanese VST

The bilingual Japanese version of the VST has been available for several years. However, in recent use of the test and subsequent review of the results we have identified several areas in need of revision. Below, we describe each of these and explain how we modified the instrument to address the issue. Two native speakers of Japanese with specializations in applied linguistics principally informed the collaborative process.

English Loanwords

Ten percent of the tested words in the first eight frequency bands of the VST exist as loanwords in Japanese (e.g., yogurt, palette). This is normal and unproblematic in itself. However, in the answer choices for such words, presenting the correct option in katakana may enable examinees to correctly respond through phonological matching alone. McLean, Hogg, and Kramer (2014) express a concern in this regard, suggesting that the presence of loanwords as answer choices might inflate estimates of vocabulary size. Addressing this problem by simply replacing katakana with paraphrases of the target word in only the correct answers might allow testwise examinees to eliminate distractors which use katakana. In our revision of the test, we therefore discarded katakana English loanwords entirely from all answer choices.

When we replaced katakana answer choices, an alternative Japanese synonym was sometimes used. For instance, *paretto* (palette) was replaced with *enogu-ita*, which might not be used widely in daily Japanese language, but is easily understood and is clearly not a phonological representation of the target English word *palette*. Alternatively, we sometimes employed a brief description of the tested word. For example, in the case of *olive*, we changed the direct translation *oribu* into *shokuyo-no-mi*, meaning *edible fruit*.

Along with English-based loanwords, there were many words written using katakana that are not related to their English equivalent. For instance, following a common practice in Japanese language, many of the names of plants and animals were written in katakana (e.g., *kani* for *crab*, *rakuda* for *camel*). We decided that such use of katakana would not be a problem as they are not phonological representations of English words and do not offer any clues when choosing the correct answer. This is consistent with recommendations by Elgort (2013) on loanwords in a bilingual Russian version of the VST.

Though we were able to replace all problematic correct answers represented in katakana words using one of the two approaches mentioned above, for a small number of distractors neither approach seemed possible because the loanwords in question were very widely used in Japanese and could not easily be expressed in other ways. In such cases, we chose to change the distractor itself. For instance, *swimming pool* could only be translated to the katakana forms *suimingu puuru* or *puuru*, both phonologically derived from English, and explaining it as "a place to swim" seemed awkward. We therefore employed the distractor *yuuenchi* (amusement park) instead. The original and edited item described is shown below:

Original VST item

pub: They went to the pub.

   a. 酒場 (a drinking place)

   b. 銀行 (bank)

   c. 商店街 (shopping promenade)

   d. 水泳プール (swimming pool)

Edited VST item

pub: They went to the pub.

   a. 酒場 (a drinking place)

   b. 銀行 (bank)

   c. 百貨店 (shopping promenade)

   d. 遊園地 (amusement park)

Similarly, the distractor *supai* (spy) could only be translated to *chohouin,* which is not so commonly used in daily Japanese and seemed problematic to be on the test. This distractor was therefore replaced by *kenkyusha* (researcher), which is another occupation that matches the other options in this item.

Part of Speech and Inflections

In the original bilingual version of the VST, the four answer choices of a small number of test items

differed in part of speech or grammatical function. This can be problematic for test takers who attempt to slot each answer choice into the example sentence provided in the question stem (Elgort, 2013). We revised such items so that all answer choices shared the same part of speech or grammatical function. For instance, an item from the second 1000 words was translated as follows in the original Japanese VST:

upset: I am **upset**.
  a. 疲れた (tired)
  b. 有名な (famous)
  c. 金持ちの (rich)
  d. うろたえた (upset)

Among these options, *b* and *c* are in a form that needs to be followed by a noun (e.g., a famous person, a rich family). In contrast, although they are in adjective forms, options *a* and *d* could also be interpreted as the past tense of being tired and being upset, and do not necessarily have to be followed by a noun. This difference could possibly make options *a* and *d* look more suitable to replace "upset" in the example sentence "I am upset." In order to eliminate such inconsistency, we revised this item as follows:

upset: I am **upset**.
  a. 疲れている (tired)
  b. 有名だ (famous)
  c. 金持ちだ (rich)
  d. 怒っている (upset)

All of these forms can come at the end of a Japanese sentence, and they look equally attractive following the phrase "I am."

Length of Answer Choices

Following Elgort (2013), who recommends the balanced presence of either single-word or descriptive-sentence distractors, we revised several items so that (a) all answer choices were short Japanese translation equivalents, (b) all choices were short phrases in Japanese, or (c) two choices were single words and two were phrases. We made sure that no items had three of one answer style and just one of the other. For instance, in the original Japanese VST, the item for *poor* appeared as follows:

poor: We are **poor**.
  a. 貧しい (poor)
  b. 幸せ (happy)
  c. とても興味がある (have a strong interest)
  d. 一生懸命には働きたくない (not willing to work hard)

In this item, option *d* is substantially longer than the others. We revised this item as shown below, to have four options of about the same length, each of which would sit naturally in place of *poor* in the example sentence:

poor: We are **poor**.
  a. 貧乏だ (poor)
  b. 幸せだ (happy)
  c. 興味がある (interested)
  d. 怠け者だ (lazy)

Changes to Stems

For two items, the stems were problematic because when translated directly into Japanese, they did not provide a natural context for the answer choices. In the item for *candid*, the example sentence provided in the stem was "Please be candid." With the word "please" in the stem, it is natural to have Japanese forms that sound like a request. However, with adjectives, it seemed very awkward to express such nuance in the answer choices. As a consequence, we chose to change the stem to "He *is* candid." This way, having basic adjective forms such as *kouhei-da* (fair) or *socchokuda* (candid) as answer choices would not pose any problem. Similarly, we changed "He felt lonesome" into "He *was* lonesome" to make the use of basic adjective forms more natural.

Conclusion

We have described an effort to improve the Japanese bilingual version of the VST. We intend to follow the revisions reported here with two further analyses. First, we plan to compare the original and edited versions of the test, giving particular attention to the effect of replacing katakana in some of the answer choices. Second, we will collect initial evidence for validation of the instrument for use with Japanese university students.

As the test remains closely based on the much-studied original VST, it can be used for purposes of research to estimate vocabulary size or vocabulary growth of Japanese learners of English, and include the same errors of being based on a now dated British corpora. The VST should not be mis-

taken as a levels test, which is intended to diagnose gaps in learners' vocabularies and may be of greater pedagogical benefit for some teachers. McLean and Kramer (2015) explain the relative advantages of size and levels tests. The authors set out not to make a new size instrument, but improve the existing Japanese bilingual VST. Thus, this paper provides and explains the criterion by which the first eight 1,000-word frequency bands of the VST were revised. The revised version of the Japanese bilingual Vocabulary Size Test described here is available from the authors' Academia.edu pages.

References

Bandura, A. (1977). *Social learning theory*. Englewood Cliffs, NJ: Prentice Hall.

Beglar, D. (2010). A Rasch-based validation of the Vocabulary Size Test. *Language Testing, 27*, 101-118. http://dx.doi.org/10.1177/0265532209340194

Day, R. R., & Bamford, J. (1998). *Extensive reading in the second language classroom*. Cambridge, UK: Cambridge University Press.

Elgort, I. (2013). Effects of L1 definitions and cognate status of test items on the Vocabulary Size Test. *Language Testing, 30*, 253-272. http://dx.doi.org/10.1177/0265532212459028

Karami, H. (2012). The development and validation of a bilingual version of the vocabulary size test. *RELC Journal, 43*(1), 53-67. http://dx.doi.org/10.1177/0033688212439359

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.

McLean, S., & Kramer, B. (2015). The creation of a new Vocabulary Levels Test. *Shiken, 19*(2), 1-11.

McLean, S., Kramer, B., Stewart, J. (2015). An empirical examination of the effect of guessing on vocabulary size test scores. *Vocabulary Learning and Instruction, 4*(1), 1-10. http://dx.dof.org/10.7820/ vli.v04.1.mclean.et.al

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

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

Nguyen, L. T. C., & Nation, I. S. P. (2011). A bilingual vocabulary size test of English for Vietnamese learners. *RELC Journal, 42*(1), 86-99. http://dx.doi.org/10.1177/0033688210390264

West, M. (1953). *A general service list of English words*. London, UK: Longman, Green, & Co.

Stuart McLean is an instructor at Osaka Jogakuin. He holds an MSc Ed (TESOL) and a PGCE, and is an Applied Linguistics doctoral student at Kansai University. He has published in *Reading in a Foreign Language*, *Vocabulary Learning and Instruction*, *Language Teaching Research*, *TESOL Quarterly,* and *Language Assessment Quarterly.*

Tomoko Ishii received her PhD from the University of Nottingham in 2005, and has been teaching English at several colleges. She is currently teaching at Meiji Gakuin University in Kanagawa, Tokyo, and is an active member of the JALT Vocabulary SIG.

Tim Stoeckel is an associate professor at the University of Niigata Prefecture in Niigata. His research is in the areas of vocabulary assessment and vocabulary growth among EFL college learners.

Phil Bennett is an assistant professor in the School of International Liberal Arts at Miyazaki International College where he coordinates the language program and teaches both academic English and team-taught liberal arts courses. He is a PhD candidate at the University of Birmingham and his research interests cover all aspects of lexical acquisition.

Yuko Matsumoto is an English instructor at Miyazaki International College. Her main research theme is the L1 use in the L2 classroom focusing on reading instruction. She graduated from Monterey Institute of International Studies (MA TESOL) and is currently a PhD candidate at Tsuda College.