様　式　Ｃ－１９、Ｆ－１９－１、Ｚ－１９　（共通）

科学研究費助成事業　　研究成果報告書

平成　３０　年　　６　月　２７　日現在

機関番号：　３４４４２
研究種目：　若手研究(B)
研究期間：　2016～2017
課題番号：　１６Ｋ１６８９０

研究課題名（和文）The creation and validation of productive and receptive vocabulary levels tests

研究課題名（英文）The creation and validation of productive and receptive vocabulary levels tests

研究代表者

Ｍｃｌｅａｎ　Ｓｔｕａｒｔ（McLean, Stuart）

大阪女学院大学・国際・英語学部・専任講師

研究者番号：１０６２４７９４

交付決定額（研究期間全体）：（直接経費）　　1,400,000円

研究成果の概要（和文）：科研費を受け行なった本研究で（１）日本人EFL学習者の受容・産出言語使用に最も適切で一般的な語換算単位はレマあるいはフレマである、（２）二言語テストは日本人大学生の受容・産出語彙知識をより正確に測定する、（３）30項目は1,000語バンドを示すのに不十分であり、的確且つ実用的である為には65項目が必要であることが証明された。これらの結果からform-recallとmeaning-recall形式のオンラインテストを作成した。このテストは受容・発表語彙に使用可能な1,000-8,000語バンドの語彙知識を測定するもので、正解・不正解がわかるようになっており、不正解の項目は正答が提示される。

研究成果の概要（英文）：This KAKEN has helped establish that (a) the flemma or lemma and not the word family is the most appropriate general word counting unit for Japanese university students in regard to both their receptive and productive language use, (b) that bilingual Japanese and English tests more accurately measure Japanese university students receptive and productive vocabulary knowledge, (c) that 30 items fail to accurately represent a 1,000-word band, and 65 items is a pragmatic balance between accuracy and practicality, and (d) that meaning-recall items better represent the construct of reading than multiple-choice items.

A self-marking online form-recall and meaning-recall tests (https://www.apps4efl.com/assignments/guest.php?type=nvlt), based on the above research findings, tests were created. Tests can be receptive or productive, test knowledge of any or all of the first eight 1,000-word bands. The test provides test administrators with dichotomously marked data and actual answers.

研究分野：Vocabulary measurment

キーワード：vocabulary　measurement　reading　receptive　productive　levels

１．研究開始当初の背景

Appropriateness of the word-family unit when measuring productive and receptive lexical knowledge

L1 studies (Nagy, Anderson, Schommer, Scott, & Stallman, 1989) posit that until the age of 18 learners develop their L1 knowledge of derived forms. L2 studies (Mochizuki & Aizawa, 2000; Schmitt & Meara, 1997; Sasao & Webb, 2017; Schmitt & Zimmerman, 2001; Ward & Chuenjundaeng, 2009) with predominantly Japanese learners of English, provide evidence that using the word-family over-estimates EFL learners' lexical knowledge. However, these studies did not directly address the appropriateness of the word-family unit with EFL learners. Despite this the word-family unit is used in corpus derived word lists ((e.g., British National Corpus (BNC) (Nation, 2006), Corpus Of Contemporary American-English (COCA), BNC/COCA), vocabulary tests based upon them ((e.g., VST (Nation & Beglar, 2007), VST bilingual L1 variants (Elgort, 2013) LVLT (McLean, Kramer, & Beglar, 2015), P-Lex (Meara & Bell, 2001), Lex30 (Meara & Fitzpatrick, 2000), A-Lex (Milton & Hopkins, 2006), NVLT (McLean & Kramer, 2015) and CATTS, (Laufer & Goldstein, 2004)), and numerous studies utilizing these corpora, lists, and tests. Thus, the validity of different lexical units to EFL learners of varying lexical proficiencies should be comprehensively investigated.

２．研究の目的

The goal of this study is to investigate Japanese EFL learners' ability to comprehend (produce the L1 meaning from the L2 form) inflectional and derivational WF6 forms consisting of a single affix or multiple affixes, and presents the following three research questions.

RQ1 Is there a significant difference in Japanese EFL learners' ability to comprehend the base form, inflectional forms, and derivational forms of the same word family?

３．研究の方法

Participants

The participants (N = 279), aged between 18 and 22, were Japanese EFL university students who had studied English formally for six years in secondary school. 185 participants (87 females and 98 males) were first-year university students enrolled in three hours per week of English language courses. They belonged to their institution's intermediate English stream. 50 participants (4 females and 46 males) were first-year university students enrolled in three hours per week of English language courses. They belonged to their institution's advanced English stream. 21 participants (16 females and five males) were second-year students who had elected to study in upper advanced English classes, and were enrolled in 3.0-4.5 hours of English language courses per week. 23 participants, ten third-year (nine females and one male) and 13 fourth-year (12 females and one male) students, were English majors. A further 19 students failed to complete both instruments used in this study, thus their data were not included. The participants came from intact classes taught by the researcher and they gave written consent for the data in this research to be used.

The majority of the participants had not completed a standardized English proficiency test. However, the participants'

New Vocabulary Levels Tests scores (NVLT) (McLean and Kramer 2015) provide an indication of the participants' written receptive meaning-recognition knowledge by which they were separated into three groups. The beginner group (n = 85) consists of participants who failed to score 22 or more out of 24, on the second 1,000-word band of the NVLT. The intermediate group (n = 177) scored 22 or more on each of the second and third 1,000-word bands, but failed to score 22 or more on the fourth 1,000-word band. The advanced group (n = 17) scored 22 or more on the fourth and/or the fifth 1,000 bands. While a cut off score of 24 out of 24 would have been preferable for mastery, such a strict cut off score would not allow for measurement error. A one one-way ANOVA with groups as the independent variable and NVLT scores as the dependent variable found that the NVLT scores among these three groups were significantly different, $F(2, 276) = 126.253$, $p = .000$. With significant differences (p = .000) in NVLT scores between the advanced group and the intermediate group, as well as the intermediate group and the beginner group. Only the advanced-group participants had all completed a standardized test, and had shared learning experiences. They were all third or fourth-year English majors who had spent a year studying abroad in English-medium classes at universities in native English speaking nations, and possessed English proficiencies very rarely met even at graduate school level in Japan, with a mean TOEIC score of 872. These participants had correctly answered at least 22 out of 24 multiple choice vocabulary base word items from the 4,000 and/or 5,000 BNC/COCA WF6 bands. Webb and Sasao (2013) refer to the first 5,000 words of the BNC/COCA the greatest range in vocabulary learning for the majority of L2 learners.

Instruments

Inflectional and derivational forms comprehension test. This study sets out to measure the participants' ability to comprehend the inflectional and derivational forms of known base words. Selecting target base forms too difficult for the participants would decrease the amount of data available for analysis. Thus, 17 low-proficiency participants' knowledge of 20 high-frequency words from the first 2,000 base words of the BNC was investigated, and twelve words (*use, move, collect, center, teach, accept, maintain, develop, standard, circle, adjust, publish*) were known by all participants. The 12 words possess too many inflectional and derivational forms to be tested in their entirety.

The test did not include inflectional forms marking plurals (*-s*) and the third person (*-s, -es*) because Japanese does not mark these forms, making the testing of them in a translation test problematic, and because they have minimal impact on reading comprehension.

The resulting 100-item comprehension (L2 form to L1 meaning translation) test (see supplementary materials) presented the target words in a context sentence using vocabulary from the first 2,000 WF6s of the BNC/COCA, or other words (including loanwords) that piloting indicated that the participants would comprehend (*panda, vocabulary, text, snack, network, appropriate*). Here is an example of one item: *The computer is now* <u>*usable*</u> = _____. Instructions for the tests were presented in the participants' L1, Japanese. Instructions accompanying the test were as follows: *Please translate the underlined word*

into Japanese. The answer can be expressed as a single word or multiple words. Please pay attention to the form of the underlined word. Please ensure that your writing is easy to read.

The items belonging to a single WF6 were presented in sets. However, the order in which target WF6 members were presented differed to prevent the participants guessing the meaning of the target form based on their order. For example, the past tense form did not always follow the base word (see Appendix). The test was completed within 30 minutes in class and was administered by the participants' teacher and author of this paper.

Separating students into lexical-proficiency groups. The NVLT (McLean and Kramer 2015) consists of 24 multiple-choice items per 1,000-BNC/COCA WF6 band for the first five 1,000-word bands. The NVLT was administered by the author one week after the inflectional and derivational forms comprehension test, and the participants completed the NVLT within 30 minutes. Examinees select the Japanese response with the closest meaning to the target word from four options. A bilingual Japanese and not monolingual English version of the NVLT was given as bilingual multiple-choice tests increase the unidimensionality of data collected and are completed more quickly (McLean *et al*. 2016). The test instructions were written in the examinees' native language (Japanese) to support rapid and accurate comprehension. Examinees were to skip questions they had absolutely no idea of. An example item is shown below. The translations in parentheses were not visible to examinees.

School: This is a big **school**.

1.

a. **銀行** (bank)

b. **海の動物** (sea animal)

c. **学校** (school)

d. **家** (house)

４．研究成果

Research question 1 asked if there is a significant difference in Japanese EFL learners' ability to comprehend the base form, inflectional forms, and derivational forms. Considering the very high threshold of 98% knowledge of tokens within a text necessary for reading comprehension, the overestimation of learners' ability to comprehend inflectional and derivational forms is a greater concern than its underestimation. Thus, for WF6 to be recommended as an appropriate general counting unit, a significant difference between participants' ability to comprehend the base form, inflectional and derivational forms should not be present. The absence of a significant difference would suggest that the participants' ability to comprehend the base word is similar to that of the tested inflectional and derivational forms of that base word, and would support the assumption made when using WF6, that comprehension of the base word is evidence of being able to comprehend all other inflectional and derivational forms belonging to that WF6. This is why when creating WF6 word lists the presence of deviational forms, for example reusable, is added to the frequency of the use WF6 in that corpus. Likewise, test-takers are accredited with knowledge of unseeingly, if they correctly demonstrate knowledge of see on the WF6-based VST. Or when the lexical load of a text is estimated with a WF6-based lexical profiler, the derived forms of centralization and

centric among other derivational forms of center are counted as high-frequency words from the first 1,000 WF6. In contrast, if a significant difference is present it indicates that the participants' ability to comprehend a base word is significantly different to that of tested inflectional and derivational forms, and would not support the assumption made when using WF6.

The Cochran's Q analysis indicated a significant difference ($p < .001$) among the number of correct responses to all 12 base words, and correct responses to associated inflectional and derivational forms of the word *use*. The large effect sizes ($\eta^2 =$. 36 - .71) of these differences are presented in Table 1. This pattern of significant differences among the number of correct responses to the base word and correct responses to inflectional and derivational forms of the same WF6 is present across all 12 of the target word families and among the beginner- ($\eta^2 =$. 44 - .71) (Table 2), intermediate- ($\eta^2 =$. 34 - .73) (Table 3), advanced- ($\eta^2 =$. 28 - .84) (Table 4) group participants, with the exception of the advanced-group participants' performance on the *publish* WF6 forms ($p = .001$, $\eta^2 =$. 19). The statistically significant discrepancies between participants' ability to comprehend the base word and other members of the same WF6 accompanied by large effect sizes (Table 1, Table 2) strongly indicate that WF6 overestimates the participants' ability to comprehend inflectional and derivational forms. Thus, considering the high threshold of 98% knowledge of tokens within a text necessary for unassisted reading comprehension, WF6 is not the most appropriate general written-receptive word counting unit for Japanese EFL learners.

Table 1

The significance and effect size of differences in the number of participants who comprehend the base form and the number of participants who comprehend associated WF6 members.

WF6	All participants
Use	$N = 278$, $p = .000$, $\eta^2 = .39$
Move	$N = 278$, $p = .000$, $\eta^2 = .71$
Collect	$N = 278$, $p = .000$, $\eta^2 = .55$
Center	$N = 273$, $p = .000$, $\eta^2 = .36$
Teach	$N = 279$, $p = .000$, $\eta^2 = .48$
Accept	$N = 255$, $p = .000$, $\eta^2 = .47$
Maintain	$N = 268$, $p = .000$, $\eta^2 = .55$
Develop	$N = 268$, $p = .000$, $\eta^2 = .41$
Standard	$N = 267$, $p = .000$, $\eta^2 = .46$
Circle	$N = 273$, $p = .000$, $\eta^2 = .68$
Adjust	$N = 263$, $p = .000$, $\eta^2 = .55$
Publish	$N = 269$, $p = .000$, $\eta^2 = .46$

＜引用文献＞

Bauer, L., & Nation, P. (1993). Word families. *International journal of Lexicography*, *6*(4), 253-279.F

Elgort, I. (2013). Effects of L1 definitions and cognate status of test items on the Vocabulary Size Test. *Language Testing*, *30*(2), 253-272.

Gyllstad, H., Vilkaitė, L., & Schmitt, N. (2015). Assessing vocabulary size through multiple-choice formats: Issues with guessing and sampling rates. *ITL-International Journal of Applied Linguistics*, *166*(2), 278-306.

Laufer, B., & Goldstein, Z. (2004). Testing vocabulary knowledge: Size, strength, and computer adaptiveness. *Language learning*, *54*(3), 399-436.

McLean, S., & Kramer, B. (2015). The creation of a new vocabulary levels test. *Shiken*, *19*(2), 1-11.

McLean, S., Kramer, B., & Beglar, D. (2015). The creation and validation of a listening vocabulary levels test. *Language Teaching Research*, *19*(6), 741-760.

Meara, P., & Bell, H. (2001). P_Lex: A simple and effective way of describing the lexical characteristics of short L2 texts.

Meara, P., & Fitzpatrick, T. (2000). Lex30: An improved method of assessing productive vocabulary in an L2. *System*, *28*(1), 19-30.

Milton, J., & Hopkins, N. (2006). Comparing phonological and orthographic vocabulary size: Do vocabulary tests underestimate the knowledge of some learners?. *Canadian Modern Language Review*, *63*(1), 127-147.

Mochizuki, M., & Aizawa, K. (2000). An affix acquisition order for EFL learners: An exploratory study. *System*, *28*(2), 291-304.

Nagy, W., Anderson, R. C., Schommer, M., Scott, J. A., & Stallman, A. C. (1989). Morphological families in the internal lexicon. *Reading Research Quarterly*, 262-282.

Nation, I. (2006). How large a vocabulary is needed for reading and listening?. *Canadian Modern Language Review*, *63*(1), 59-82.

Nation, I. S. P., & Beglar, D. (2007). A Vocabulary Size Test. *The Language Teacher, 31* (7), 9-13.

Sasao, Y., & Webb, S. (2017). The word part levels test. *Language Teaching Research*, *21*(1), 12-30.

Schmitt, N., & Meara, P. (1997). Researching vocabulary through a word knowledge framework: Word associations and verbal suffixes. *Studies in second language acquisition*, *19*(1), 17-36.

Schmitt, N., & Zimmerman, C. B. (2002). Derivative word forms: What do learners know?. *Tesol Quarterly*, *36*(2), 145-171.

Ward, J., & Chuenjundaeng, J. (2009). Suffix knowledge: Acquisition and applications. *System*, *37*(3), 461-469.

Webb, S. A., & Sasao, Y. (2013). New directions in vocabulary testing. *RELC Journal*, *44*(3), 263-277.

５．主な発表論文等

（研究代表者には下線）

〔雑誌論文〕（計 1 件）

① McLean, S. (2017). Evidence for the Adoption of the Flemma as an Appropriate Word Counting Unit. *Applied Linguistics*, amw050, https://doi.org/10.1093/applin/amw050

〔学会発表〕（計 2 件）

1. **McLean, S**., & Kramer. (2017). The Word Family or the Lemma? Evidence for the Adoption of the Lemma with Japanese EFL University Students. American Association for Applied Linguistics (AAAL) 2017 Conference in Portland, Oregon, USA.
2. **McLean, S**., Nation, P., Pinchbeck, G., Brown, D., & Kramer, B. (2016). Revisiting the Word Family: What is an Appropriate Lexical Unit for Japanese EFL Learners? Vocab@tokyo, Tokyo.

〔図書〕（計 0 件）

〔産業財産権〕

○出願状況（計 0 件）

○取得状況（計 0 件）

６．研究組織

(1)研究代表者
MCLEAN, Stuart
大阪女学院大学・国際・英語学部・専任講師
研究者番号：１０６２４７９４