Development of a New Vocabulary Component in English Communication 2: MGU Vocabulary Booster

Tomoko Ishii, Dax Thomas

Background

The importance of vocabulary in second language (L2) learning and use was established decades ago, and there is abundant research on its influence on L2 performance. For instance, Qian (2002), Laufer and Aviad–Levitzky (2012), and many other studies have shown how the knowledge of vocabulary affects the extent to which learners can understand the passage they are exposed to. Research on text coverage (i.e. the percentage of running words that are covered by the words a reader can recognize the meaning of) has shown that for reasonable comprehension of a written passage, as high as 95-98% of the running words need to be recognized (e.g. Nation, 2006). The link between vocabulary and aural skills has also been shown. Stær (2009) examines the contribution of vocabulary knowledge to listening comprehension, and claims that it is responsible for half of the variance occurring in the study, and Alharthi (2020) demonstrates how the knowledge of the first 3000 words affects the learners' speaking ability.

Echoing the well recognized importance of vocabulary knowledge, Japanese students preparing for university entrance exams make efforts to expand their vocabulary, as shown in the large selection of vocabulary textbooks we can find in bookstores. Unfortunately, however, unless they enter the university with a great interest in further developing their English skills, most of them do not spend time to maintain the vocabulary they must have acquired in the course of preparing themselves for the exam. There is some research showing how quickly Japanese university students can lose their vocabulary after admission (e.g. Okamoto, 2007), and systematic support is required to maintain and reinforce the vocabulary knowledge they already have.

Another problem is that the way vocabulary is learned through vocabulary textbooks is often very superficial, focusing mostly on translation or Japanese-English definition matching. Some students seem to think that knowing the primary meaning of a word in Japanese is enough to understand the word. However, as Nation (2013) suggests, there are numerous aspects involved in "knowing" a particular vocabulary item; in order to use a word effectively, learners need to know not only its meaning but also its grammatical behavior and any specific connotations or nuances it might have. As a consequence of exam preparation focusing only on meaning recognition, the vocabulary knowledge possessed by Japanese university students often lacks depth, and this can lead to various issues

when they come to use those words in production. For instance, the words *peculiar* and *unique* share the same translation 独特な (dokutokuna), and some students say *peculiar* when they simply mean *unique* as the translation fails to convey the slightly negative connotation that the word *peculiar* has. Also, we often find clumsy constructions in student writing because their learning was focused mainly on individual words, rather than on how each word is actually used in context. For instance, some students might say *dense coffee* when they mean *strong coffee*. Grammar features can be a problem too, and some students do not seem to pay much attention to a word's part of speech information. This can lead to a grammatically wrong sentence such as *I'm sorry for absenting yesterday*.

Given the recognised importance of vocabulary knowledge, the often-observed attrition of vocabulary in students' lexicons, as well as various problems regarding vocabulary use, a project was launched to create an 'MGU Vocabulary Booster' (MVB) component for mandatory English Communication 2 courses offered by the Center for Liberal Arts at Meiji Gakuin University. The objective of developing this component is to boost the vocabulary knowledge of the students by helping them do the following three things:

1. Maintain their knowledge of basic vocabulary that supports communication in English.
2. Elaborate on their understanding of words they are likely to have studied prior to college admission.
3. Gain basic Data-Driven Learning (DDL) skills to foster their inductive vocabulary learning both during the course and after completing the mandatory EFL program.

This paper provides background to the MVB so that it may be better understood by the teachers in charge of English Communication 2 courses. It is also hoped that this paper will encourage the academic and constructive discussion on how the MVB could be amended and expanded in the future. The following sections will describe how the MVB was designed, with particular focus on the choice of word list, its presentation to the students, and the actual activities given to the students.

Designing the MGU Vocabulary Booster
Choice of word list

As is stated in the first of the three objectives, the MVB aims to maintain the knowledge of high frequency words, rather than to attain a large vocabulary size by teaching low frequency words. This is because high frequency words are heavily used in both spoken and written communication, and the knowledge of the most frequent 3,000 words is essential to comprehension of spoken and written discourse (Schmitt & Schmitt, 2014). Furthermore, the increase in communication ability gained from the study of lower-frequency words is not commensurate with the effort required to learn them. As the

target students of the MVB are non-English majors, rather than aiming to attain a larger vocabulary size, we decided it would be more beneficial for them to consolidate their knowledge of the words that support the basis of communication skills.

The understanding that some words are more useful to learners than others was established long ago, and one of the first major systematic attempts to compile a list of useful vocabulary is West's (1953) *General Service List*. With the increasingly easy access to computers of recent decades, a number of attempts have been made to produce more up-to-date lists based on large corpus data. The lists based on the British National Corpus (BNC) and the Corpus of Contemporary American English (COCA) are among the currently most influential ones. For the new vocabulary component, we decided to use the New General Service List (NGSL; http://www.newgeneralservicelist.org) for the following reasons.

First, the NGSL is a publicly available frequency-based word list, and it has abundant accompanying learning and assessment instruments that are also freely available online. With these, we can easily administer pre- and post-tests to track student progress, as well as encourage students to study the list outside of the classroom. The way the NGSL counts 'a word' was also determined suitable for our students. Some lists count a head word, its inflections and some of its derivations (e.g., *act, acts, action, actor, acting*) as a word, which corresponds to Level 6 of Word Family (Bauer & Nation, 1993). However, 'a word' in the NGSL takes only a head word and its inflected forms, which is often called *lemma* or *flemma* in the field (Pinchbeck, 2014). McLean's (2018) study on Japanese university students revealed the inappropriacy of using Word Family as a word counting unit because they do not possess enough derivation knowledge, and the use of flemma rather than Word Family made the NGSL a reasonable choice for our project.

The creators of the NGSL claim its 2801 words cover 90% of the vocabulary found in general English corpora. Although the objective of the MVB component is to consolidate students' knowledge of high frequency words, some of the words in NGSL are extremely basic. For instance, it includes function words such as pronouns and auxiliaries, which are not worth keeping in the list. Some content words (e.g. dog, have) were also selected for exclusion, and from the list of 2801 words, 511 words that met any of the following criteria were discarded:

1. Function words (articles, pronouns, prepositions, conjunctions, and auxiliaries)
2. Items in the first 500-word level on the New JACET 8000 words list (JACET, 2016)
3. Interjections (e.g. hi, hello)

The remaining 2290 words were broken into 46 sets, 50 in each of the first 45 sets, and 40 in the last. These were then uploaded to Quizlet.com. Quizlet.com is an online flashcard learning system

and was chosen as our main platform not only because it is easy to manage and freely available, but also because students do not need to create an account in order to access the vocabulary lists they are asked to learn. The site also hosts optional supplementary learning activities and games for each flashcard stack.

Activities were then prepared for each set, in three sections, each section having its own different objective. These activities can be presented in an A4 size handout, or imported into an online learning management system. The design and the objective of the activities are explained below.

Activities
a) Form-meaning matching

The first section, form-meaning matching, asks students to find a matching item for each of the three different English definitions given (Figure 1). The primary purpose of this activity is to have students go through all 50 words in the set, so that they can refresh their memory of the words they are likely to know. Should they come across a word they have no prior knowledge of, they are encouraged to take notes.

Another objective of this section is to familiarize students with dictionary-style English definitions. This is because the Japanese translation presented in dictionaries and wordbooks very often fails to capture the entire picture of what each word means, and the habit of learning words only through L1 translation, as noted above with the *peculiar* vs. *unique* example, seems to cause problems in students' vocabulary knowledge. By seeing the definitions in English for the words they are likely to have learned by high school, students are encouraged to notice the difference between the word as they know it in L1 translation and the word as they see it in the English definition. Also, by providing students with the opportunity to get used to reading and understanding English definitions, it is hoped that in the future, they will be more likely to actively choose to search for information about a word using English.

1) 50 個の単語を全て見て、次の定義に当てはまる単語を探しましょう。その際、知らない単語があったらメモをしておきましょう。	
a level of study that is completed by a student	
for a particular reason, purpose, etc.	
something that happens as a result, consequence	

Figure 1. Sample activity: Form-meaning matching

b) Data-driven learning

The second section gives students an opportunity to practice some DDL skills. DDL is an inductive approach to teaching how vocabulary items are used. In DDL learners engage in their own linguistic search by accessing corpus data directly (Szudarski, 2018). By using corpora to explore how words are used, learners can derive more information than can be found in teaching materials such as dictionaries and grammar books (Chambers, 2010). It helps students recognise the complexity of how vocabulary items are used and reinforces the idea that simply knowing the L1 translation of a word is not the same as knowing how to use that word.

Researchers emphasize the importance of knowing the lexical patterns, or collocations, in which a word is used (e.g. Lewis, 2000). Various studies show that active attempts should be made to raise learners' awareness in this area and cultivate such knowledge (e.g. Ying & O'Neill, 2009, Boers & Lindstromberg, 2012); DDL is aptly suited to this. Also, it can help learners see the differences between two words that often share a common Japanese translation word, as well as identify how English-based loanwords can be different in meaning or usage from their original English words.

Although DDL is an approach with many benefits, learners need to be trained to use corpora effectively. This section therefore is designed to have them learn how to use, and what to look for in, a typical general English corpus. From among the freely available corpus tools, we selected SkELL (http://skell.sketchengine.co.uk/run.cgi/skell) for the current project because, being specifically designed for language learners as opposed to researchers, it has a user-friendly interface and a simplified set of tools. It does not allow us to run the types of advanced search that corpus researchers often perform, but it still gives sample sentences as well as collocation information, which is sufficient for DDL purposes.

The initial search page of SkELL is quite simple (Figure 2). We simply type in the word we want to learn about (for example, *communicate*), and choose one of the three search types. *Examples* gives us 40 example sentences, from which we can see various aspects of the word such as grammar patterns and common meanings. *Word sketch* gives us a summary of the word's usage in terms of its common collocates. In the case of a verb, for example, it shows the typical subjects, objects, modifying adverbs, as well as other verbs that are used with the conjunction *and* (e.g. *communicate and collaborate, communicate and interact*). The final type of search, *Similar words*, shows other related words that are often used in a similar context as the searched word, and is useful when looking for synonyms and semantically-related alternatives.
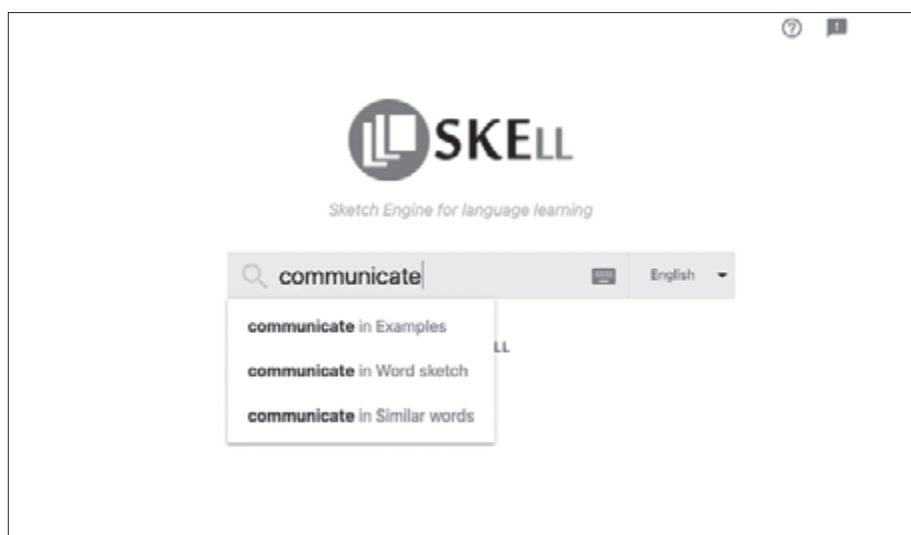
Figure 2. SkELL search interface

In this section, for a small number of words from each set, learners are asked to perform one of several different search-related tasks on SkELL (Figure 3). These tasks include: finding common collocates for a verb or an adjective, finding differences between two or more similar words (e.g. *dramatically* vs. *greatly*), exploring the use of a word when it occurs in a less well-known part of speech form (e.g. *fuel* as a verb), and exploring differences between English words and how their katakana counterparts are used in Japanese (e.g. *highlight* vs. *hairaito*).

These activities are designed to encourage students to pay closer attention to how words are actually used. It is also expected that students learn there are patterns in the way each word is used, and those patterns can be different even when two or more words seem similar in meaning. Finally, it is hoped that, eventually, students will actively use these resources when they use English on their own trying to reach a higher level of sophistication.

2) fuel（燃料）という単語には動詞としての使い方もあります。fuel の動詞について SkELL を使って知識を深めてみましょう。

どんな意味？	どんな主語？	どんな目的語？	その他気づいたこと、面白いと思ったこと

調べたことを意識して、fuel（動詞）を使ったオリジナル例文を作ってみましょう。

___

Figure 3. Sample activity: DDL

c) Derivation and part of speech exercise

The final section is for learning various derivation forms, and aims to address some of the vocabulary knowledge weaknesses of Japanese learners identified in previous research. McLean (2018) demonstrates how Japanese university students lack knowledge of derivations, and that this is certainly a knowledge they need to expand. Having good derivation knowledge would mean that when they learn a new word in the future, they will be able to anticipate its derived forms, thus adding not just one word to their lexicon, but several.

Additionally, Ishii et. al (in press) describe the difficulties Japanese learners have when a word form occurs in more than one part of speech. For example, the learner may know the meaning of the noun form *silence*, but not necessarily understand what its verb form means in the sentence, "The voices of the laborers were silenced." The knowledge that a single word form can be used in different parts of speech can help learners parse the word when seeing it take on an unfamiliar grammatical role in a sentence.

To cultivate students' awareness of parts of speech and knowledge of various derivation forms, this section asks students to find derivation forms of given words, and then fill in the blank in sentences with a word form that is appropriate both in meaning and form (Figure 4).

3) 単語は品詞によって形が変化したり、同じ形が２つ以上の品詞で使われたりすることがあります。以下の表を埋め、品詞に注意しながら下の例文の空欄の意味に合う単語で埋めましょう。

	名詞	動詞	形容詞	副詞
例）honest	honesty	×	honest	honestly
silence				
thick				
inspire				

Dr. Richard gave us a very _____ presentation.

It took me many days to finish this _____ book.

Not knowing what to say, we kept eating _____.

Figure 4. Sample activity: Derivation

Conclusion

This paper reported on the development of a new vocabulary component, the MVB, for English Communication 2 classes. It is based on a 2290-word list created using the NGSL, and solid knowledge of those words should provide a good basis for communication in English. As covering all these words in one academic year would not be feasible, teachers can select the level of words they wish their students to learn and focus on those in class. Even if all the exercises cannot be directly presented during class time, all the materials are available for self-study outside of the class.

Previous research findings suggest the tendency of Japanese learners to lose their vocabulary after entering the university, and indicate problems stemming from a rather superficial understanding of the words they study. The MVB, therefore, focuses on aiding the retention of the words that are the most useful for communication and deepening the knowledge of the words they have already learned.

At the time of writing this paper, the MVB has not been presented to students yet, and whether or not this boosts our students' vocabulary knowledge and usage has yet to be tested. Especially, the potential contribution of DDL to the expansion of their vocabulary knowledge as well as their learning habits is worth closer attention. We hope that the MGU Vocabulary Booster component will lead

learners to a better command of the basic vocabulary items that serve as a foundation for both written and spoken communication.

References

Alharthi, T. (2020). Investigating the relationship between vocabulary knowledge and FL speaking performance. *International Journal of English Linguistics, 10*, 37–46.

Boers, F., & Lindstromberg, S. (2012). Experimental and intervention studies on formulaic sequences in a second language. *Annual Review of Applied Linguistics, 32*, 83–110.

Browne, C., Culligan, B., & Phillips, J. (2013). *The New General Service List*. Retrieved from http://www.newgeneralservicelist.org.

Chambers, A. (2010). What is data-driven learning? In A. O'Keeffe & M. McCarthy (eds.) *The Routledge Handbook of Corpus Linguistics*. London and New York: Routledge, 356–358.

Ishii, T., Bennett, P., & Stoeckel, T. (in press). Challenges in the assumptions of using a flemma-based word counting unit. *Vocabulary Learning and Instruction*.

JACET (2016). *Daigaku Eigo Kyouiku Gakkai Kihongo Risuto Shin JACET 8000. (Japan Association of College English Teaching List of Basic Words: New JACET 8000)*. Tokyo: Kirihara.

Laufer, B. & Aviad–Levitzky, T. (2012). What type of vocabulary knowledge predicts reading comprehension: Word meaning recall or word meaning recognition? *The Modern Language Journal, 101*, 729–741.

Lewis, M. (Ed.). (2000). *Teaching collocation: Further developments in lexical approach*. Hove, UK: Language Teaching Publications.

McLean, S. (2018). Evidence for the adoption of the flemma as an appropriate word counting unit. *Applied Linguistics, 39*, 823–45.

Nation, I. S. P. (2006). How a large vocabulary size is needed for reading and listening. *Canadian Modern Language Review*, 63, 59–82.

Nation, I. S. P. (2013). *Learning vocabulary in another language*. Cambridge, England: Cambridge University Press.

Okamoto, M. (2007). Lexical attrition in Japanese university students: A case study. *JACET Journal, 44*, 71–84.

Pinchbeck, G. G. (2014, March 22–25). Lexical frequency profiling of a large sample of Canadian high school diploma exam expository writing: L1 and L2 academic English [Roundtable presentation]. *American Association for Applied Linguistics Annual Conference*, Portland, Oregon.

Stær, L. S. (2009). Vocabulary knowledge and advanced listening comprehension in English as a foreign language. *Studies in Second Language Acquisition, 31*, 577–607.

Szudarski, P. (2018). *Corpus Linguistics for Vocabulary: A guide for research*. London and New York: Rout-

ledge.

Qian, D. (2002). Investigating the relationship between vocabulary knowledge and academic reading performance: An assessment perspective. *Language Learning*, *52*, 513–536.

West, M. (1953). *General Service List of EnglishWords*. Boston: Addison-Wesley Longman.

Ying, Y., & O'Neill, M. (2009). Collocation learning through an 'AWARE' approach: Learner perspectives and learning process. In A. Barfield & H. Gyllstad (Eds.), *Researching collocations in another language* (pp. 181–193). Basingstoke, UK: Palgrave Macmillan.