The Global Academic Vocabulary Lexicon: A New ELT Resource

Paul Wadden[1]
International Christian University

Dan Ferreira[2]
International Christian University

Edward Rush[3]
Trinity College, University of Melbourne

Abstract

This article describes the Global Academic Vocabulary (GAV) lexicon, lessons, and platform that was initially implemented at International Christian University in Tokyo and is now under significant further development at the University of Melbourne and NYU-Tokyo. Research by Na and Nation has shown that understanding of about 95% of the words in an academic text is required for learners to confidently comprehend its meaning. But exactly what words do university learners need to know to achieve such a level of coverage? The GAV provides one important answer to this question by combining the headwords from the three most significant long-standing corpus-based vocabulary studies to date: the University Word List (UWL), the Academic Word List (AWL), and the English for Academic Purposes (EAP) word lists, with a fourth, the New Academic Word List (NAWL), now being added. This article provides the rationale behind the creation of the GAV.

Issue

> *With a vocabulary size of 2,000 words, a learner knows 80% of the words in a text which means that one word in every five (approximately two words in every line) are unknown. Research by Na and Nation (1985) has shown that this ratio of unknown to known words is not sufficient to allow reasonably successful guessing of the meaning of the unknown words. At least 95% coverage is needed for that.*
> (Nation & Waring, 1997)

This key observation by Nation and Waring (1997) sets forth what should be the axiomatic principle for EAP vocabulary study: attaining a level of word knowledge that permits the learner to proficiently (if not effortlessly) read typical university texts and to generally comprehend (if not completely absorb) academic lectures in English. Words are indeed "polysemous"—as any linguist or lexicographer knows—yet by acquiring the

---

[1] Paul Wadden is a senior lecturer at International Christian University, Mitaka City, Tokyo, Japan. Correspondence should be sent to paulwadden@gmail.com

[2] Dan Ferreira is an instructor at International Christian University, Mitaka City, Tokyo, Japan. Correspondence should be sent to ferreira@icu.ac.jp

[3] Edward Rush is an instructor at Trinity College, University of Melbourne, Melbourne, Australia. Correspondence should be sent to ferreira@icu.ac.jp erush@trinity.unimelb.edu.au

primary meanings of core academic vocabulary learners can reach a level of proficiency in which they are more likely to succeed in their university study in English, whether in English-medium classes in Japanese universities or in college courses in overseas universities. This is why Coxhead contends in her seminal article introducing the AWL in *TESOL Quarterly* in 2000 that "An academic word list should play a crucial role in setting vocabulary goals for language courses, guiding learners in their independent study, and informing course and material designers in selecting texts and developing learning activities" (Coxhead, 2000, p. 214). In addition, Cummings (1994), among many others, has further observed that academic vocabulary contributes across the board to the development of Cognitive Academic Language Proficiency (CALP) that enables language learners to apply language in their university courses.

But the question remains: exactly what words should ELT students learn to reach the "95% coverage" level?

Initial Answers

Corpus linguistics has yielded definitive—yet imperfect—answers to this question. Empirical analyses of word frequency (millions of pages from academic texts on several continents) have already identified the core academic vocabulary that learners need to know to comprehend, with a reasonable level of confidence, typical university-level academic texts and lectures. Here, in brief, are the figures and findings.

> 836 headwords, University Word List (UWL) (Xue & Nation, 1984) (a word list created by combining four pre-existing lists, two based on corpora and two based upon frequently annotated words by students in textbooks)
>
> 874 headwords, EAP core word list (EAP) (Masuko, Mizoguchi, Sano, Shiima, Thrasher, & Yoshioka, 1997) (based upon English texts in use at a Japanese liberal arts college)
>
> 570 headwords, Academic Word list (AWL) (Coxhead, 2000) (based upon a large corpus of texts written for an international audience but mainly sourced in New Zealand)
>
> 963 headwords, New Academic Word List (NAWL) (Browne, Culligan, & Phillips, 2013) (based upon the Cambridge English Corpus) [NAWL headwords are now being added to the GAV]

These word lists—from different geographic regions and academic contexts—have inspired teachers, textbook writers, and more recently, software makers, to create programs to empower EFL/ESL learners to reach the *Holy Grail* of 95% coverage. Yet none of the lists, in itself, will get students to that ultimate goal. The reason is that with fewer than 1,000 words in each list, none provides sufficient coverage of frequently appearing academic vocabulary.

A New Solution

The Global Academic Vocabulary (GAV) lexicon and Learner's Dictionary were created to meet this challenge as well as to embrace the global diversity of academic English. Initially constructed at International Christian University, Tokyo (Wadden, 2013), the GAV is now being further developed and refined by EAP teachers at the University of Melbourne and at New York University's American Language Institute in Tokyo. To help

students reach the crucial 95 percent threshold, the GAV combines the headwords from the three most important traditional corpus-based vocabulary studies mentioned previously, and introduces them to students in 17 lessons that progress from the most common to the least common words (by relative frequency count). A weekly quiz for each lesson (or half lesson) motivates students to study the words, provides a basis for classroom assessment, and gives students feedback on their progress.

The pedagogical premise of the GAV is efficacy: to lower as much as possible the "learning burden" (Nation, 2006, p. 70), and the labor it takes for a learner to learn the words. To achieve this, the GAV Lessons and Learner's Dictionary (1) identify the essential core academic vocabulary (e.g. the verb "access"), (2) pinpoint the important related words to learn at the same time (i.e., the noun "access" and the adjective "accessible"), (3) provide the primary meaning of the words in simple English as well as (for Japanese learners of English) in their bilingual Japanese form, (4) draw attention to the common phrases the words appear in (such as "access the internet"), and (5) demonstrate their "use in context" in sample sentences.

The GAV Dictionary is posted online and can also be downloaded in a searchable PDF file; two GAV-related websites created on Quizlet further offer—for each GAV lesson—flashcards, sample sentences, auto-recordings, bilingual exercises, and quizzes for students' self-study. At the University of Melbourne, the GAV is currently being administered through the Schoology platform; at ICU, Edmodo and paper-only versions have been used.

To promote the generative use of the vocabulary from the lists, teachers encourage students to consolidate and extend their use of the words in classwork and written assignments, such as by including GAV words in essays and highlighting them in different font color for quick identification and salience. Vocabulary Profilers such as AntWordProfiler (a free downloadable app) can also identify for students the headwords that appear in any specific online or e-reading material to reinforce their appearance and underscore how valuable it is to learn them.

Student Response and Evaluation

Since 2009 when the GAV was first introduced to one of the three tracks of students in International Christian University's English Program, the response of students has been overwhelmingly favorable. For instance, in the highest track, one sample of 80 students (ITP TOEFL scores of 530-650) gave these assessments:

> 96% evaluated the Academic Reading and Writing (ARW) GAV vocabulary program as being from "somewhat" to "very useful" (78% "useful" or "very useful"; 18% "somewhat useful"; while only 3% regarded the GAV program as "not useful"; with 1 percent no answer)

A sample of 80 students from the mid-level track (ITP from 480-530) appraised the GAV program as follows: *86% of students rated the vocabulary program as "good" or "very good" (13% as "poor" or "very poor")*

In 2015, Trinity College of the University of Melbourne prepared more than 800 international students for entry to the university's degree programs. The GAV was applied on a limited basis to observe whether it would assist in increasing students' vocabulary knowledge. Results were promising, with students in two cohorts showing in just ten weeks an average 20% increase in vocabulary recognition and recall after undertaking the program.

Students entered the program by taking the first version of Paul Nation's 100-item Vocabulary Size Test. They then undertook independent study in preparation for GAV progress tests using preparatory exercises hosted in the Quizlet flashcard application. At the end of the program, students were retested using the second version of Nation's Vocabulary Size Test. Students in the lower-level cohort (IELTS 6-7) achieved an average score improvement of 20.0% (entry average: 11816.7 words; exit 13816.7 words), while students in the upper-level cohort (IELTS 6.5-9) achieved an average score improvement of 19.6% (entry average: 13116.7 words; exit 15600 words). About 400 Trinity College students are now using the GAV.

To date, the GAV lessons and lexicon as used in classrooms and independent study online comprise only the EAP, UWL, and AWL word lists; at present, six more lessons covering 469 additional words and 13 prefixes from the NAWL are being added to the lessons, lexicon and dictionary.
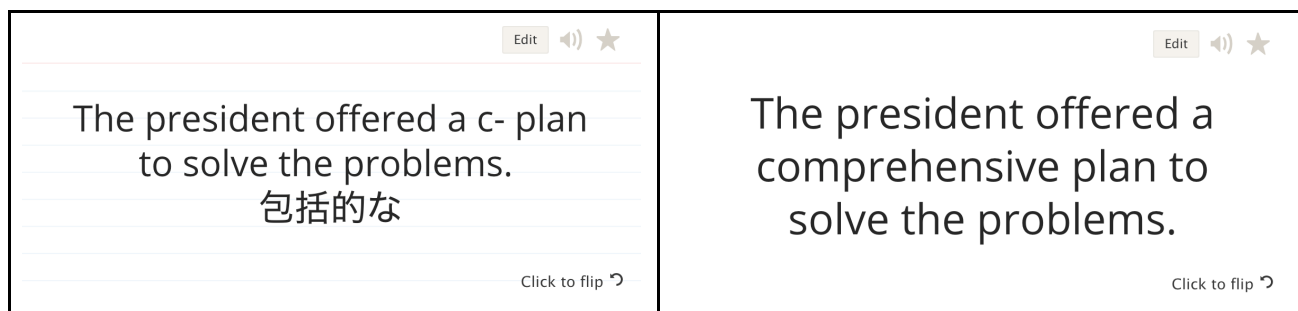
Quizlet

Given the prevalence of smartphones in the classrooms, the vocabulary app called Quizlet ("Quizlet," n.d.) was used to create electronic word decks of each of the lessons of the GAV. (Search for "GAV" from the Quizlet website to obtain the decks.) Although there are vocabulary learning apps (such as "Anki") that are more customizable and have superior scheduling algorithms based on the difficulty of the word, the prices for the apps are prohibitively expensive for students. Quizlet, on the other hand, is free and is available on the two most popular mobile operating systems: Android and iOS. Despite its limitations, Quizlet has many attractive features on both the online platform and the mobile apps. For example, once the decks are downloaded onto a smartphone, the user can edit each card or deck, adding or removing information as needed.

The Decks' Design

For each of the Quizlet decks, one sample sentence per vocabulary item in each of the lessons was extracted from the GAV Dictionary (Wadden, 2013). Figure 1 is an example of the design of a card. The front facing card contains a sentence in the target language; the target word is identified as the initial letter and a hyphen. The sentence is followed by a gloss in L1. In most cases, a single L1 gloss is provided as a prompt; however, there are times when more than one L1 gloss is needed to provide additional context.

Figure 1
A screenshot from Quizlet (online) of a vocabulary card from GAV deck 11.
Note: the left image is the "front" and the right image is the "back"

The decision to present the vocabulary items in the form of sentences is based on research that suggests learning new words in context can help learners to develop deeper, more elaborate semantic connections in their mental lexicon (cf., Laufer & Shmueli, 1997; Nation, 2001; Nation & Webb, 2011; Oxford & Crookall, 1990). Connecting target words to a context of other familiar words provides a "cognitive foothold" that increases memory performance and a reliable trace to meaning and understanding on a receptive level (Laufer & Shmueli, 1997, p. 91). In the process of guessing from context, especially with verbs, there is the added challenge of guessing grammatically correct patterns (such as the past tense of regular vs. irregular verbs). Learners may begin to notice natural chunks of language that collocate with the target word. Sternberg & Powell's research (as found in Nation & Webb, 2011, p. 78) has found that guessing from context contributes to other skills such as reading and speaking; thus, going beyond the mere acquisition of target words. To reinforce pronunciation and listening practice, the user has the option of switching on the "voice" feature on Quizlet. Some students have reported using this feature with their headphones while studying the vocabulary of the week on their commute to and from school.

Another noteworthy feature of the card is the use of priming by providing the initial letter of the target word. This serves two purposes: (1) to reduce the recall burden on the learner and (2) to foster an orthographic relationship as the learner tries to decode from the clues provided. Research suggests that the technique of providing an initial-letter cloze slows the attrition rate. Successful guessing of the cloze can also provide a satisfactory accomplishment and ensure a semantic knowledge critical to recall (Burton, Niles, & Wildman, 1981, p. 162). Similar to Nation's idea of using stems as primers for learning target words (Nation, 2008, p. 110), adding a suffix primer for verbs in the third person or past tense can encourage grammatical accuracy in guessing. Students can be further advised to edit the card and remove the primer should they find guessing the correct target word too easy.

Although the long-held practice of L2 to L1 vocabulary learning remains popular among Japanese students today, Rivers and Temperley argue that the direct translation method could result in being a "crutch" that inhibits a deeper processing (as found in Laufer & Shmueli, 1997, p. 93). In a study that examined the effectiveness of several primers—L1 translation; L2 definitions; L2 synonyms; and pictures—Laufer and Shmueli found L1 translation to be the most effective overall, especially for beginning learners (Laufer & Shmueli, 1997). It should be noted that the Rivers and Temperley study did not include contextualized sentences, rather the results focused on the attrition of words learned out of context. Moreover, given the range of the words in the GAV, students may

encounter words that they have never seen before. As suggested earlier, the students could be directed to delete the L1 gloss or replace it with a synonym if they need the added challenge. Note that the back-facing card, in Figure 1, does not include the L1 gloss, thus reducing the dependence on the "crutch." Admittedly, providing an initial-letter cloze as a primer and an L1 translation straddles the line between receptive and productive retrieval. Overall, it is hoped that learning the target words in this way will encourage the generative approach to learning vocabulary. Supplementing the GAV Quizlet decks with exploratory writing practices (e.g. using dictionaries or an online corpus) could foster a process that encourages better retention of the target words.

Classroom Practice

Two of the researchers in this study have the privilege of conducting classes with their students three times a week, which is rare in the Japanese university context. Meeting the students on such a regular basis is an opportunity to reinforce the practice of structured review and spaced learning. At the start of the term, students are briefly introduced to Pimsleur's scale as a guide for reviewing vocabulary (See Nation, 2001, p. 77). Essentially, the newer the target word is, the quicker the attrition rate. To foster the practice of ongoing review, be it the vocabulary from the week before or the deck for the current week, an overhead projector can display a deck from an app on a mobile device, as the students are getting ready for the start of class. To discourage "serial learning" (Nation, 2008; Thornbury, 2006), Quizlet can be set to display the cards in random order. Furthermore, to model and encourage participation among those students who are already settled in their seats, the teacher can participate in the guessing of target words; praising students who may have guessed quicker or by expressing satisfaction in one's correct guesses. Pausing the display to "star" a card that is difficult to learn or by turning on the "voice" feature to model pronunciation are a couple of ways to encourage students to be active learners of the vocabulary using the app. Using the last five minutes of class to display Quizlet visual illustrations is also a routine to not only signal the close of the class but also to reinforce the need to include spaced learning.

Quizzes

Edmodo is a free learning management system that can be used to create vocabulary quizzes ("Edmodo," n.d.). In addition, should one decide to create multiple-choice options, the questions from the quizzes can be randomized, a useful feature that can reduce the level of anxiety for students working in proximity to one another (e.g., working side-by-side on a quiz in a computer room). Moreover, Edmodo Quizzes can be scheduled ahead of time. One of the most useful features of Edmodo is how it keeps track of all the answers for each test. Students can be encouraged, for example, to take note of the question items they answered incorrectly and "star" cards in their Quizlet app for further, more intense review. From the teacher's perspective, one interface can be used to see how an individual student performed, and another interface offers the ability to see the whole class performance on the test. More specifically, the application provides pie charts for each question indicating what percentage of the class answered correctly or incorrectly. For a review quiz, the instructor can easily "cherry-pick" the most common incorrect

answers and use the same questions for a final quiz; thus providing some incentive, in this case grade-based, for students to review their vocabulary.

GAV Development, Further Implementation, Vocabulary in Japan

Quizzes and flashcards, including English-only and bilingual English-Japanese versions, are now available to teachers and learners on Quizlet websites and the Schoology and Edmodo platforms, and the GAV materials have been converted into Blackboard format and into GIFT format so they can be loaded directly into Moodle.

Teachers who wish to use the GAV can also request access to tests for each lesson (these are not available on websites so as to preserve their assessment value); these tests, one for each half-lesson, can motivate students to study the words, give them feedback on their progress, and provide a basis for assessment.

At present, words from the watershed New Academic Word List (NAWL) by Browne, Culligan and Phillips (2013), an extensive new list showcased at the 2014 JALT Vocabulary Sig Symposium (2014) and one of the most important and exciting new English vocabulary resources in the world, are being added to GAV to expand its coverage. Yet many questions remain.

How can Japanese EFL students in particular, and English-language learning students around the world in general, best be provided with the core vocabulary that will power and empower their language learning? Can national textbook publishers—and Ministries of Education—be inspired to systematically present foundational vocabulary to a nation's students during primary and secondary school education (such as the core vocabulary of the GSL or NGSL) which can then be followed by focus on the GAV or NAWL during higher education? What platforms and pedagogies can best promote students' optimal vocabulary learning, both in the classroom and independently? These are but a few of the pressing questions—and promising solutions—to be explored.

References

Browne, C., Culligan, B., & Phillips, J. (2013). New General Service List. Retrieved June 28, 2015, from http://www.newgeneralservicelist.org/

Burton, J. K., Niles, J. A., & Wildman, T. M. (1981). Levels of processing effects on the immediate and delayed recall of prose. *Journal of Literacy Research, 13*(2), 157-164.

Coxhead, A. (2000). A new academic word list. *TESOL Quarterly, 34*(2), 213-238. doi:10.2307/3587951

Cummins, J. (1994). The acquisition of English as a second language. In K. Spangenberg-Urbschat & R. H. Pritchard (Eds.), *Kids come in all languages: Reading instruction for ESL students* (pp. 36-62). Newark, DE: International Reading Association.

Edmodo. (n.d.) Edmodo. Retrieved from https://www.edmodo.com/, accessed January 1, 2016.

Laufer, B. (1989). What percentage of text-lexis is essential for comprehension? In C. Laurén & M. Nordman (Eds.), *Special language: From humans thinking to thinking machines* (pp. 316-323). Clevedon: Multilingual Matters.

Laufer, B., & Shmueli, K. (1997). Memorizing new words: Does teaching have anything to do with it? *RELC Journal, 28*(1), 89-108.

Masuko, M., S. Mizoguchi, M. Sano, K. Shiina, R. Thrasher, and M. Yoshioka. 1997. *English vocabulary for academic purposes: Analysis of data from the physical and social sciences*. Tokyo: Liber Press.

Na, L., & Nation, I. S. (1985). Factors affecting guessing vocabulary in context. *RELC Journal, 16*(1), 33-42. doi:10.1177/003368828501600103

Nation, I. S., & Waring, R. (1997). Vocabulary size, text coverage and word lists. In N. Schmitt & M. McCarthy (Eds.), *Vocabulary: Description, acquisition and pedagogy*. Cambridge: Cambridge University Press. Retrieved from http://www.fltr.ucl.ac.be/fltr/germ/etan/bibs/vocab/cup.html

Nation, I. S. (2001). *Learning vocabulary in another language*. Cambridge: Cambridge University Press.

Nation, I. S. (2006). How large a vocabulary is needed for reading and listening? *The Canadian Modern Language Review / La Revue Canadienne Des Langues Vivantes, 63*(1), 59-81. doi:10.1353/cml.2006.0049

Nation, I. S. P. (2008). *Teaching vocabulary: Strategies and techniques*. Boston: Heinle.

Nation, I. S. P., & Webb, S. (2011). *Researching and analyzing vocabulary* Boston: Heinle, Cengage Learning.

Oxford, R., & Crookall, D. (1990). Vocabulary learning: A critical analysis of techniques. *TESL Canada Journal, 7*(2), 9.

Quizlet. (n.d.) Quizlet. Retrieved from https://quizlet.com/, accessed December 31, 2015.

Thornbury, S. (2006). *How to teach vocabulary*. Malaysia: Pearson Education.

Wadden, P. (2013). Vocabulary: GAV and GAV Dictionary. Retrieved from https://sites.google.com/site/paulwaddenforstudents/vocabulary

Xue, G., & Nation, I. S. P. (1984). (1984). A university word list. *Language Learning and Communication, 3*(2), 215-229.