TEACHING ENGLISH FOR CHEMISTRY AT A JAPANESE UNIVERSITY

Makoto Shimizu*, Masaki Murata**, and Kris Ramonda*

*Faculty of Science, Tokyo University of Science

**Graduate School of Engineering, Tottori University

makoto@rs.kagu.tus.ac.jp

Abstract: Although studying English presents considerable difficulties for Japanese university students, in this time of globalism, our students are expected to read and write academic articles in English. To address this problem, we chose a basic word list of 2188 words, and compiled an academic word list of 776 words for a sub-discipline of organic chemistry using a "tailored" corpus. The two lists combined cover 85.72% of words in 100 papers from an academic journal, which would facilitate students' access to the sub-discipline in English.

Keywords: ESAP, academic word lists, tailored corpora

Introduction

Reading and writing academic papers in English presents considerable difficulties for Japanese university students. One important contributing factor to this difficulty is the students' lack of appropriate vocabulary needed to consume and produce academic papers in English. In a Japanese context, these lexical gaps can be attributed, in part, to the absence of informed and carefully created word lists based on students' needs. In order to begin approaching the word coverage necessary to read academic articles in English, not only do learners need to know high frequency words from a basic word list, but also a sufficient number of academic words connected to their specific sub-discipline. Thus, one means of alleviating this difficulty that Japanese students face is by employing corpora in the selection and compilation of basic and academic word lists.

As for a basic word list, although the Japanese Ministry of Education and Science states that a student should learn 3,000 English words at junior and senior high school, it does not specify exactly which words to learn. Thus, an English instructor at a Japanese university needs to select which basic word list to provide students with prior to teaching academic words. However, there are some reasons why basic word lists compiled in English-speaking countries such as *GSL*, *NGSL*, *the BNC/COCOA lists* are not appropriate as the basic word lists in Japan. Another issue is an academic word list: we need a concise and precise word list for our students to learn efficiently over a short period of time. To solve this problem, Coxhead (2000a, b) proposed The Academic Word List, hereafter AWL for short. She claimed, together with a basic word list of 2,000 words, her AWL covers 95% of any kind of academic writing. Since then, a number of English for General Academic

Purposes (hereafter EGAP) academic word lists have been compiled. Hyland & Tse (2008), however, who contend that it is impossible to compile a list for academic English in general. Instead, they posit that all we can do is compile academic word lists for each individual discipline, namely, English for Specific Academic Purposes (hereafter ESAP) academic word lists. Since then, several ESAP academic word lists have been compiled. In this research, we discuss issues concerning a basic word list, an academic word list for Chemistry, and how to compile an academic word list for Chemistry using a "tailored" corpus.

Materials and Methods

We chose JACET2188 as the basic word list for two reasons: 1) the list is based on English teaching textbooks published by major Japanese textbook companies and 2) the list is oriented towards academic writings, which differentiates itself from lists complied in English-speaking countries (e.g. GSL, NGSL, BNC, COCA, Oxford2000/3000), which are oriented towards conversation. We designed our academic word list to be ESAP for chemistry majors. Moreover, we decided it should not be chemistry in general but a sub-discipline of chemistry, namely, organic chemistry. This would allow for us to develop more fine-tuned word lists for those learners involved in this sub-discipline of chemistry. Furthermore, we selected one academic journal of organic chemistry rather than several academic journals of organic chemistry. It is crucial that the list is entirely made for the journal. This reason is that a highly customized list for a specific journal would allow for learners to engage in the narrow reading of highly similar articles. As learners begin to learn words from the word lists, not only will their word coverage increase for the journal's articles, but they will also be able to deepen their lexical knowledge through seeing the same words many times and in varied contexts. By developing a list in such a way, we "tailor" it to suit the specific needs of a learner population, which in this case are Japanese chemistry students who engage in the narrow reading of organic chemistry in a targeted journal. That is why we call the list "tailored". Just like tailors pick up a fabric, we picked up one particular academic journal. To the best of our knowledge, this is the first academic word list targeted for a single academic journal. 100 papers from an online journal, *Journal of American Chemical Society* (hereafter JACS for short) were chosen, around 0.6 million in total. We turned the 100 papers into an online corpus database, which allowed us to retrieve words automatically. Out of the corpus, we retrieved words automatically. See Shimizu & Murata (2010, 2012, 2013, & 2015) for how to compile a corpus and retrieve words. After compiling the corpus, we parsed the texts, tagged and stemmed words, extracted all words according to occurrences, and obtained nouns, verbs, adjectives, adverbs, and others. Following this, we selected frequent words, which we defined as those words occurring over 50 times for nouns, verbs, and adverbs, and over 100 times for adjectives and others. Next, we deleted words listed in New JACET2188 from the academic word list unless they were: 1) listed with different parts of speech or, 2) technical terms. The final stage was also key to this research because frequency searches alone cannot account for the full range of what makes a word useful. Thus, we asked chemistry experts for their opinions on English for the journal, and deleted and added words accordingly.

Results and Discussion

As a result, an academic word list of 598 nouns, 134 verbs, 38 adjectives, 4 adverbs, 4 others, and 787 words in total, were obtained.

First, we will turn our attention to the nouns. The top 10 most frequent nouns are *Figure* (1064 times), *datum* (742), *protein* (722), *molecule* (495), *complex* (459), *substrate* (480), *pH* (438), *NMR* (426), *spectra* (425), and *mechanism* (419). We might regard a few words such as *Figure*, *datum* and *mechanism* as EGAP. Observe *Figure* is spelt with the capital *F* here because the word is used as *in Figure 3.2*. EGAP words, however, consists of just 14.16% of all nouns in the list. The majority of nouns are ESAP words such as *protein, molecule, complex, substrate, pH, NMR,* and *spectra*. Note the word *complex* here is not an adjective but a noun, "a type of compound."

The top 10 most frequent verbs are *observe* (595), *support* (519), *determine* (425), *contain* (419), *bind* (415), *suggest* (405), *form* (378), *obtain* (377), *provide* (364), and *indicate* (335). The majority of the top 10 most frequent verbs are EGAP words such as *observe, support, suggest, obtain, provide, indicate, determine* and *contain*. However, we also have ESAP words such as *bind* and *form*. When we consider the whole verbs, we recognise that EGAP words are only 28.57 % of all verbs of the list.

EGAP words such as *experimental (480), binding (362), molecular (346), structural (298), catalytic (269), solvent (251), significant (223), consistent (211), available (205), radical (179)* are the top 10 most frequent adjectives. However, we also have ESAP words such as *binding, molecular, catalytic,* and *solvent.* EGAP words consist of 37.57% of adjectives in the list. The top 10 most frequent adverbs, *respectively* (199), *previously* (164), *significantly* (131), *approximately* (100), *relatively* (97), *typically* (67), *furthermore* (67), *readily* (64), *experimentally* (64), *negatively* (61) are all EGAP words, but we also have ESAP words. EGAP words consist of 28.57% of adverbs in the list. Others, mainly prepositions and conjunctions, *due to* (101), *such as* (86), *whereas* (84), and *as well as* (54), are all EGAP words.

Let us now discuss whether the academic word list should be EGAP or ESAP. According to Paquot (2010), her Academic Keyword List (henceforth AKL), an EGAP academic word list consisting of 354 nouns, 233 verbs, 180 adjectives, 87 adverbs, and 75 others, covers 95% of any academic writing, including academic journal articles, textbooks, PhD dissertations, MA theses, and lab reports. We compared AKL with The JACS academic word list. It turned out that only 18.29% of all words are common. The nouns common to AKL and The JACS academic word list, for instance, are *Figure, datum,* and *mechanism,* which are general, but occur frequently in the academic journal.

In contrast, nouns occurring only in the JACS academic word list are specific to Chemistry. Note that the word *complex*, for example, is not an adjective meaning "difficult to understand" but a noun meaning "a kind of compound". We took chemistry experts' advice and deleted quite a few words such as *sugar* and *H*, which they considered too basic. There was another set of words we deleted after following their advice such as *alamethicin*, *bicelles*, and *dysfanction*, which were too technical. We also added words which chemistry professors recommended. In other words, their opinion aided in fine-tuning the list for maximum utility.

We now turn to nouns occurring only in AKL, namely, *ability*, *action*, *advice*, *adult*, *age*, and so on. We see that they are not relevant to Chemistry. We could safely claim that an EGAP list might not be very efficient.

We checked the coverage of New JACET2188 and the JACS academic word list. We discovered that the coverage of the former is 70%. Incidentally, the coverage of General Service List is 63%, which suggests that the JACET list is a more suitable basic list for students. The coverage of the JACS academic word list is 15.72%. The total coverage is 85.72%, and we would like to state that it is satisfactory for a start.

Conclusion

To conclude, we claim that a basic word list and an ESAP academic word list can provide an efficient platform for English for Chemistry majors reading academic papers in English. We would like to improve the coverage of 85.72% to 90% in the near future, and 95% ultimately, so that English learners can reach the threshold necessary for reading seamlessly without the need to consult a dictionary.

Acknowledgements

A part of this research was supported by the Grants-in-Aid for Scientific Research of Japan Society for the Promotion of Science (No 24520655). We would like to thank Professors Koichi Tsukiyama, Motoyuki Simonaka, Hidetoshi Kawai, and Kazuo Miyamura at Tokyo University of Science for their valuable comments on English for Chemistry. Without their cooperation, this project has been unattainable.

References

Makoto Shimizu (Academic Word Lists)

http://www.rs.kagu.tus.ac.jp/makoto

Shinichi Ishikawa (NewJACET8000)

http://language.sakura.ne.jp/s/voc.html

Charniak, Eugene. (2000). A maximum-entropy-inspired parser. *Proceedings of NAACL*, 132-139.

Coxhead, A. (2000a). The Academic Word List: A Corpus-based Word List for Academic Purposes. in

Coxhead, A. (2000b). A New Academic Word List. *TESOL Quarterly* 34:2, 213-238.

Hyland, K. & Tse, P. (2007). Is there an "Academic Vocabulary"? *TESOL Quarterly* 41:2, 235-253.

Paquot, M. (2010). *Academic Vocabulary in Learner Writing*, London: Continuum.

Schmid, H. (1997). Probabilistic part-of-speech tagging using decision trees. In D. Jones & H. Somers (eds.) *New methods in language processing*. London: UCL Press.

Shimizu, M., Murata, M., Nakatani, Y. & Hijikata, Y. (2010) Vt+N patterns in physical chemistry papers. *GloCALL 2010, Kota Kinabalu, Malaysia 1-3 December, 2010.*

Shimizu, M., & Murata (2012). Passive and active voices in scientific papers. *AsiaTEFL 2012, New Delhi, India 4-6 October, 2012*.

Shimizu, M., & Murata, M. (2013). Passive and active voices in engineering papers. *GloCALL 2013, Da Nang, Viet Nam 28-30 November, 2013*.

Shimizu, M., & Murata, M. (2015). Active and passive voices in biology. *CamTESOL 2015*, Phnom Penh, Cambodia *28 February-1 March, 2015*. http://camtesol.org/past-conferences/2015-conference/Stream EAP.zip/EAP03.

West, M. (1953). *A General Service List of English Words.* London: Longman.

