Feature Presentation

Building Essential and State-of-the-art Vocabulary in Information Science: Observations from Classes at a Japanese University

Hiroshi Kohsaka

The University of Electro-Communications

The terminology of information science has increased significantly, which makes it difficult to effectively cover all essential technical terms in a class in one given semester. To provide students with an opportunity to improve their English proficiency, an autonomous class activity to collect keywords in information science was incorporated into an English course at a Japanese university. Results suggest that students in their second year and above can build both basic and state-of-the-art vocabulary autonomously.

The vocabulary in information science has expanded at a rapid pace in recent years. Since the invention of the first programmable digital computer ENIAC (Electronic Numerical Integrator and Computer) in 1946, novel technical terms have been added to the vocabulary of information science (Comer, 2017). Furthermore, the number of technical words is increasing in association with the exponential development of new technology, including artificial intelligence (Russell & Norvig, 2022) and quantum computation (Nielsen & Chuang, 2010).

Vocabulary acquisition is one of the core elements of language learning (Hu & Nation, 2000; Thornbury, 2002). A number of teaching strategies to support building vocabulary have been developed, such as verbal interaction with others (Wang, 2015), consulting dictionaries (Zhang et al., 2021), note-taking (Jin & Webb, 2021), and guessing from context (Nassaji, 2006). These effective learning strategies are implemented based on reliable vocabulary lists built in existence (Nation & Beglar, 2007; Browne et al., 2013). Thus, establishing

relevant vocabulary lists is prerequisite to conduct vocabulary building activities in ESP (English for Specific Purposes) classes. However, in the field of information science, although several vocabulary lists of the technical terms have been published (Esteras & Fabre, 2007; Hirai, 2012), it is still challenging for university teachers and students to follow the rapid pace of the appearance of new words due to the fact that published lists can become obsolete quickly.

This paper reports a survey and the results of an English class in a Japanese university aiming at the development of vocabulary in information science by a student-centered activity. Within a theme selected by a teacher, students selected relevant keywords at their discretion. The comparison between the keywords collected by students and a vocabulary textbook demonstrates two-fold results. First, the vocabulary developed by a large majority of students successfully nominated basic keywords in information science. Second, the students' list covers state-of-the-art words related to novel computer technology. These observations suggest that an autonomous class activity would be a viable option to tie together two areas of vocabulary in information science: a) permanent and foundation vocabulary and b) new and potentially transient vocabulary.

Methods

The data for this study were collected in ESP classes taught at the University of Electro-Communications in Tokyo, Japan. The classes were elective-compulsory subjects; each student selected one class among other compulsory language classes. The students were in their second year or above and have passed both Academic Written English and Academic Spoken English classes (taught by other instructors) in their first year. Therefore, the students possess adequate command of English for the specific needs of information science. The information science classes were taught face-to-face in the standard 15-week semester, and each class meeting lasted 90 minutes once per week. The students' proficiency in English is intermediate level, and they possess high literacy in information science.

A typical instructor-supervised student-centered activity took the following form throughout the course. The instructor assigned one topic from information science in each class (Table 1).

Table 1

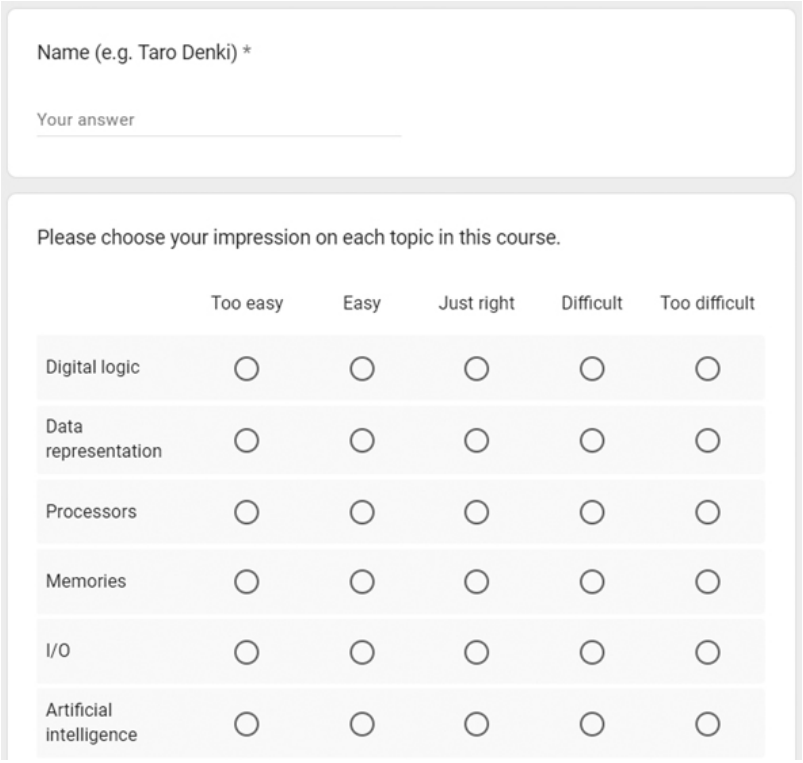
The Topic of Each Class

Week	Topic	Week	Topic
1	Discrete math	7	Artificial intelligence
2	Digital logic	8	Deep learning
3	Data representation	9	Computer vision
4	Processors	10	Robotics
5	Memories	11	Quantum computing
6	Input/Output devices		

At the beginning of the class, each student brainstormed the topic for five minutes to find a keyword about the topic. The instructor allowed the students to access the Internet during this activity. Students conducted group discussions with two to four classmates to share their keywords. At the end of the class, each student searched for one keyword in five minutes and submitted both the keyword and its explanation to Google Forms. In total, 306 keywords were collected in 11 weeks. In the last week of the course (week 15), the instructor conducted a survey with a questionnaire to ask about the rigor of the course for reference to improve next year's class, in particular, to adjust the speed and the level of the course to the students' demand. The survey included multiple-choice single-answer questions (Figure 1). Eighty out of 87 students agreed to participate in this study. The instructor received informed consent from all the students whose work was used in this report. Python 3.7 was used to perform all the statistical analyses. The questionnaire responses about the level of the course were analyzed using the one-way ANOVA (analysis of variance) and post hoc Tukey test.

Results

The 306 keywords collected in the class were compared with the vocabulary list from a published book listing the essential vocabulary in information science (*N*


Name (e.g. Taro Denki) *

Your answer

Please choose your impression on each topic in this course.

	Too easy	Easy	Just right	Difficult	Too difficult
Digital logic	○	○	○	○	○
Data representation	○	○	○	○	○
Processors	○	○	○	○	○
Memories	○	○	○	○	○
I/O	○	○	○	○	○
Artificial intelligence	○	○	○	○	○

Figure 1. Part of the questionnaire conducted to ask about the level of the course

= 512 words in Hirai, 2012). Figure 2 shows a Venn diagram of the number of overlapping words between student-selected words and the textbook vocabulary. The students' list included 27 essential keywords in information science (Figure 3).

To test whether the students could acquire state-of-the-art terminology, the instructor included four modern technology fields in the topic list ("Artificial intelligence," "Deep learning," "Computer vision," and "Quantum computing", Table 1). Students could find some state-of-the-art words through this class activity. As a consequence, several state-of-the-art keywords were listed in the vocabulary selected by students but not in the basic textbook (Figure 4): "object recognition" and "face authentication" from artificial intelligence; "quantum interference" and "q-bit" from quantum computing.

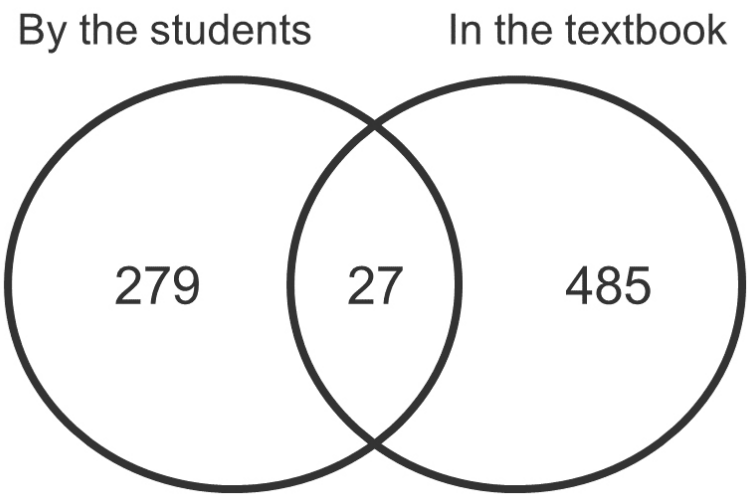


Figure 2. Comparison of the vocabulary lists created by the students and described in the textbook

algorithm	bit	input	byte	error	pixel	CPU
storage	bus	register	memory	cache	decode	volatile
scanner	flash memory	interface	monitor	mouse	smartphone	screen
hardware	speech recognition	layer	search	OCR	simulation	

Figure 3. A list of terms selected by the students overlapping the vocabulary in the textbook

This observation demonstrates that autonomous vocabulary building is capable of developing a word list including keywords from new technology in information science. One of the advantages of this strategy is that teachers do not have to prepare the comprehensive vocabulary list beforehand. In reality, it is impossible for English teachers to cover the wide range of information science fields. The student-centered activity is not only effective for students to learn vocabulary but also practical for teachers to provide keywords from new technology.

To evaluate the course level for the students, the instructor conducted a survey with a questionnaire to ask students about the difficulty of the class

algorithm	bit	input	byte	error	pixel	CPU
storage	bus	register	memory	cache	decode	volatile
scanner	flash memory	interface	monitor	mouse	smartphone	screen
hardware	speech recognition	layer	search	OCR	simulation	

Figure 4. A list of terms selected by the students that did not overlap the vocabulary in the textbook

activity (Figure 1). Students felt the class level was neither too easy nor too difficult for most topics (Table 2).

It should be noted, however, that the one-way ANOVA yielded significant variation in difficulty among topics, $F(9, 744) = 9.226$, $p < 10\text{-}12$. The post hoc Tukey test demonstrated that quantum computing was challenging for some students with a significance level at $p < 0.001$. In sum, whereas the topics and other class design, such as the progression and structure of content, should be carefully chosen, as a whole, it appears that the topics selected in the classes suited the students.

Discussion

This article has reported results from a survey in ESP classes that attempted to build a vocabulary list for information science with an autonomous class activity at a Japanese university. The list developed in classes included both essential keywords and state-of-the-art terminology. Furthermore, the results from the questionnaire show that the topics and rigor of the autonomous class activity matched the level of the students in their second year or above majoring in computer and information science. These observations indicate that building vocabulary through students' autonomous activity could constitute one learning method to keep up with the ever-growing vocabulary in information technology. This strategy will be applicable to vocabulary building in other fields evolving at a fast pace where English teachers cannot keep up with the state-of-the-art terminology.

A couple of issues remain unexplored. First, the basic vocabulary used in this article was from a book published over 10 years ago (Hirai, 2012). That age

Table 2

Results of the Questionnaire about the Level of the Course

Topic	Too easy	Easy	Just right	Difficult	Too difficult
Digital logic	2 (3%)	22 (28%)	41 (51%)	13 (16%)	2 (3%)
Data representation	2 (3%)	12 (15%)	49 (61%)	16 (20%)	1 (1%)
Processor	1 (1%)	15 (19%)	46 (58%)	18 (23%)	0 (0%)
Memories	2 (3%)	21 (26%)	46 (58%)	9 (11%)	2 (3%)
Input/Output device	4 (5%)	18 (23%)	44 (55%)	11 (14%)	3 (4%)
Artificial intelligence	2 (3%)	20 (26%)	42 (54%)	13 (17%)	1 (1%)
Deep learning	2 (3%)	7 (9%)	41 (51%)	27 (34%)	3 (4%)
Computer vision	0 (0%)	11 (14%)	45 (56%)	24 (30%)	0 (0%)
Robotics	1 (3%)	9 (24%)	21 (57%)	6 (16%)	0 (0%)
Quantum computing	0 (0%)	4 (5%)	28 (35%)	37 (47%)	10 (13%)

may cause an underestimation of the number of keywords between the students' choices and the basic vocabulary. Thus, an updated essential vocabulary list will be required for adequate evaluation of the word list. Second, although several state-of-the-art words were found in the students' list, the quality of the words needs to be verified. To implement autonomous vocabulary building in classes, it would be necessary to consider a way to verify the adequacy of the keywords in the list. One possible strategy to guarantee the quality of words would be teachers' helpful intervention so that the students can recognize what vocabulary is more important. Lastly, building vocabulary lists is not the goal and should be followed by the next step—to design effective and practical methods to use the vocabulary list aiming at the improvement of students' proficiency. The utility of the vocabulary list should be amplified when it is used with proper strategies. Adopting efficient learning strategies can enable students majoring in information science to master the requisite vocabulary in their ever-growing

field. In particular, the students' ability to update their vocabulary is undoubtedly essential after they graduate from university.

Conclusion

An autonomous class activity was implemented in ESP classes in a Japanese university to develop a functional vocabulary list in the field of information science. The list built through the classes included both essential and state-of-the-art vocabulary in information science. This pedagogical intervention will help teachers and students in design and implement effective procedures to acquire the necessary vocabulary with greater facility in the actual application outside of the classroom environment. Having an ability to learn novel vocabulary by themselves is a critical skill to survive in an ever-growing and competitive world.

References

Browne, C., Culligan, B., & Phillips, J. (2013). The new academic word list. http://www.newgeneralservicelist.org.

Comer, D. (2017). *Essentials of computer architecture* (2nd ed.). CRC Press.

Esteras, R., & Fabre, E. M. (2007). *Professional English in use ICT: For computers and the Internet.* Cambridge University Press.

Hirai, M. (2012). *Kikutan science: Information science*. Alc.

Hu, M., & Nation, P. (2000). Unknown vocabulary density and reading comprehension. *Reading in a Foreign Language*, *13*(1), 403–430.

Jin, Z., & Webb, S. (2021). Does writing words in notes contribute to vocabulary learning? *Language Teaching Research*. https://journals.sagepub.com/doi/pdf/10.1177/13621688211062184

Nassaji, H. (2006). The relationship between depth of vocabulary knowledge and L2 learners' lexical inferencing strategy use and success. *The Modern Language Journal*, *90*, 387–401. https://doi.org/10.1111/j.1540-4781.2006.00431.x

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

Nielsen, M. A., & Chuang, I. L. (2010). *Quantum computation and quantum*

information (10th ed.). Cambridge University Press.

Russell, S., & Norvig, P. (2022). *Artificial intelligence: A modern approach* (3rd ed.). Pearson.

Thornbury, S. (2002). *How to teach vocabulary*. Pearson.

Wang, K.-H. (Isobel). (2015). The use of dialogic strategy clusters for vocabulary learning by Chinese students in the UK. *System*, *51*, 51–64. https://doi.org/10.1016/j.system.2015.04.004

Zhang, S., Xu, H., & Zhang, X. (2021). The effects of dictionary use on second language vocabulary acquisition: A meta-analysis. *International Journal of Lexicography*, *34*, 1–38. https://doi.org/10.1093/ijl/ecaa010

Author bio

Hiroshi Kohsaka *is an associate professor at the University of Electro-Communications, Tokyo, Japan. He has an interest in developing materials for English for science and technology. kohsaka@uec.ac.jp*

Received: January 20, 2023
Accepted: May 18, 2023