THE JOURNAL OF ASIA TEFL Vol. 1, No. 1, pp. 333-347, Spring 2004

JACET 8000 and Asia TEFL Vocabulary Initiative

Toshihiko Uemura
Siebold University of Nagasaki, Japan

Shin'ichiro Ishikawa
Kobe University, Japan

Indigenization of English in many Asian ESL countries and partial learning of English in EFL countries are two major hindrances to Asian Englishes today. Although Englishes are growing day by day, we have not had any common ground on which we can compare our English with other English varieties. The authors of this article were organizing members of the Committee of Revising JACET (Japanese Association of College English Teachers) List of Basic Words. In this article they claim that their experiences compiling JACET word list are applicable to making word lists for Asian English users and learners, which become their vocabulary base of comparison, and the basis for the Asia TEFL Vocabulary Initiative.

In Asia, the Chinese language has the predominant share in terms of the number of native speakers. However, Yun and Jia (2003)'s 'China English' and Zhu (2003)'s 'New Challenges in ELT in China' suggest English has become the most favored language for global communication in China. This is not surprising because rapid development of information technology and the expansion of global commerce are going hand in hand with English.

Here, a question is raised: What variety (or varieties) of English is/are widely used? As Kachru's concentric circles of English clearly indicates, English is a language which has more non-native speakers than native speakers, about which Trudgill and Hannah (2002) write:

> Equally as important, we believe that native English speakers travelling to areas such as Africa or India should make the effort to improve their comprehension of the non-native variety of English (much as Americans would have to improve their comprehension of ScotEng when traveling in Scotland) rather than argue for a more English-type English in these areas. (p. 124)

In her article in *English Today*, Dr. Elizabeth Erling, who is a specialist of Global English and now teaching in Germany, expresses a more drastic view:

> In a world where English both functions as a global language and is appropriated to several different local contexts, it seems as if we are clinging to an out-dated model of a standard ideology that is no longer possible or even useful to maintain. (Erling, 2002, p. 10)

She believes neither British English nor American English can function as the Global Standard. Graddol's futurologist view in *The Future of English* supports her viewpoint:

> The indications are that English will enjoy a special position in the multilingual society of the 21st century: it will be the only language to appear in the language mix in every part of the world. … Yesterday it was the world's poor who were multilingual; tomorrow it will also be the global elite. So we must not be hypnotised by the fact that this elite will speak English: the more significant fact may be that, unlike the majority of present-day native English speakers, they will also speak at least one other language – probably more fluently and with greater cultural loyalty. (Graddol, 1997, p. 63)

Evidently, the focus of English has been changing from the 'Inner Circle' to the 'Outer and Expanding Circles,' and then to 'Asian varieties of English.'

When it comes to varieties of Englishes or World Englishes, we have been busy discussing 'indigenization of English' and 'hegemony of English,' and failed in creating any objective ground on which our own variety of English is compared with other Englishes. Now Asian ESL/ EFL users and learners should have their own criteria, which are based on neither British English nor American English. JACET 8000 is a new word list for this purpose.

WHAT IS JACET 8000?

In 2003, the Committee of Revising JACET List of Basic Words published “JACET List of 8000 Basic Words” (thereafter JACET 8000). The JACET 8000 is a radically new word list designed for all English learners in Japan. This list is based on two kinds of corpora: the British National Corpus (BNC) and JACET 8000 sub-corpus. Although the BNC consists of 100 million words, most of them are taken from British English texts that are several-decades-old, and English texts for learners are hardly included. Therefore, the committee has compiled a corpus of approximately six million words to supplement the BNC. Its data comes from the recent American newspapers, magazines, and scripts of TV program or cinema, and also from children’s literature, junior or senior high school English textbooks, and various English tests conducted in Japan. The content of the sub-corpus is shown below:

TABLE 1
The Content of JACET 8000 Subcorpus (Murata, 2003, p. 358)

Genre	Content	American / British	Spoken/ Written	Size (mil. words)
Mass media	Newspapers and Magazines	A/B	W	1
	Script of TV Programs	A/B	S	1
Cinema	Cinema Script	mainly A	S	1.3
Educational	Junior or Senior High School Textbooks	mainly A	mainly W	0.8
ESP	Exams (University Entrance Examination Test, STEP, TOEFL, TOEIC)	mainly A	W	0.2
	Scientific Articles	mainly A	W	0.3
Literature	Children Literature	mainly B	W	1.2

Two kinds of corpora naturally produce two kinds of frequency information of each word. When comparing them, the committee has adopted the statistical score of log-likelihood (See Rayson and Garside (2000) and Sugimori (2003) for detail). Leech, Rayson and Wilson (2001) show three reasons for the need to adopt log-likelihood ratio or G^2:

> 1. We need a statistic that does not require the data to be distributed in a particular pattern. Many statistical tests assume that data are in a so-called 'normal distribution.' With linguistic data such as word frequencies in texts this is often simply not the case and invalidates the use of such measures.
> 2. We need a statistic that does not over- or under- estimate the significance of a difference between two samples. The Pearson chi-squared test, one of the most commonly-used measures, has been shown to over-estimate the importance of rare events; the G^2 has been proved better in this regard.
> 3. We need a statistic that is insensitive to differences of size between two samples. Again, Pearson's chi-square test has been shown to be poor in this respect, whereas G^2 performs better. (p. 16)

The use of log-likelihood ratio, which has been rarely attempted by creators of previous major word lists, seems to give considerable reliability to the rank information of each word in JACET 8000. Table 2 shows log-likelihood scores of three sample words:

TABLE 2
Frequency and Log-likelihood Score
(Committee of Revising JACET List of Basic Words Ed., 2003, p.108)

	POS	freq in the BNC	freq in Subcorpus	log-likelihood
look	noun / verb	126930	13972	1875.1
relevant	adjective	7950	153	− 427.6
buy	verb	25582	1875	0.0

The scores above illuminate how much the rank in the sub-corpus deviates from that in the BNC. If the absolute value of the score is sufficiently large, the original rank in the BNC is appropriately modified. After conducting some educational adjustments, the committee finally chose the most important 8000 words. (For detail, see Ishikawa, 2003a).

JACET 8000 is a unique word list in many ways: First, its 8000 words were presented in eight levels (from 1 to 8) in accordance with the frequency and educational significance of each word. Second, taking English learners at their first stage into consideration, 250 introductory words were chosen as the ‘plus250,’ which supplements with the base 8000 words. Third, these word lists were released both in the paper and electronic forms. These are main features by which JACET 8000 is claimed to be the most reliable and usable English world list.

TEXT COVERAGE

There are several studies which prove how well JACET 8000 covers the vocabulary in many English texts. For example, Mochizuki, another JACET committee member, compares JACET 8000 and its former version, JACET 4000, and concludes that the former is a more reliable word list than the latter: Mochizuki (2003b) reveals that the top 2000 words in JACET 8000 cover approximately 90% of the words used in the graded readers for beginners, and its 8000 words cover approximately 95% of them. However, JACET 8000 covers only 85% of the words appearing in English newspapers.

TABLE 3
Text Coverage of JACET 8000 (Committee of Revising JACET List of Basic Words Ed., 2003, p.113; Mochizuki, 2003b, p. 54)

sample texts		~ 1000 wds	~ 2000 wds	~ 8000 wds
graded readers	Stage1: 400 wds	85.7%	88.6%	92.1%
	Stage2: 700 wds	86.3%	89.4%	93.5%
	Stage3: 1000 wds	87.7%	91.1%	93.3%
newspapers		70.3%	75.2%	85.0%

Laufer (1992) insists that a reader needs to know at least 95% of the words in a given text to understand it efficiently. From his estimate, JACET 8000 may be effective only for the graded readers, but not for the newspapers. Isn't JACET8000 a reliable word list when we read them or other authentic English texts?

Shimizu, another committee member, developed an online JACET 8000 add-on program, 'JACET 8000 Level Marker,' which analyzes the words in actual English texts, and indicates their word ranks. Following is a sample output:

> Wildfires earlier_1 this_1 year_1 stripped_3 away_1 the_1 vegetation_5 that_1 would_1 have_1 normally_2 held_1 the_1 ground_1 in_1 place_1.

In this case, the words, 'earlier,' 'this,' 'year,' 'away,' 'the,' 'that,' 'would,' 'have,' 'held,' 'ground,' and 'place' belong to the level 1 or the words ranked 1-1000; the word 'normally' to the level 2 or those ranked 2001-3000; the word 'stripped' to the level 3 or those ranked 3001-4000; and the word 'vegetation' to the level 5 or those ranked 5001-6000. Meanwhile, the word 'wildfires,' which is shown no rank information, is not included in JACET 8000.

By Shimizu's Level Marker, we examined the three samples which are three short sample passages of 100 words taken from newspapers, examinations, and governmental documents. The words in the boldface are those not included in JACET 8000. The first example is shown below:

Sample 1. A news article in The New York Times (from nytimes.com)

> With_1 hopes_1 for_1 finding_1 anyone_1 alive_2 in_1 the_1 rubble_8 fast_1 fading_3, teams_1 from_1 a_1 number_1 of_1 countries_1, including_4 the_1 United_1 States_1 and_1 **Switzerland**, scaled_2 back_1 their_1 search_1 and_1 rescue_2 efforts_1 and_1 turned_1 their_1 attention_1 to_1 the_1 thousands_1 of_1 survivors_5, many_1 of_1 them_1 suddenly_1 homeless_5 and_1 **bereaved**. With_1 the_1 surge_5 of_1 assistance_4, the_1 tiny_2 **airstrip** that_1 serves_1 this_1 ancient_2

> backwater on_1 the_1 Silk_3 Road_1 was_1 **clogged**. State_1-run_1 television_1 said_1 **16,000** people_1 had_1 been_1 buried_2 in_1 the_1 three_1 days_1 since_1 the_1 **predawn** quake in_1 what_1 had_1 once_1 been_1 a_1 city_1 of_1 more_1 than_1 **80,000**. Estimates_2 of_1 the_1 final_1 toll_6 have_1 bobbed_6 up_1….

Here, according to a mechanical calculation, the text cover rate is 93%. However, among nine unregistered words, two are numerals and one is a proper noun. This suggests that JACET 8000 covers almost all of the words in the newspapers, except for three difficult words of 'bereaved,' 'airstrip,' and 'clogged.'

The second example is taken from the Standard Aptitude Test (SAT), a commonest college admission test in the USA. The passage below is a part of the reading material in the 'Critical Reading' section:

Sample 2. A Reading material in the SAT (from collegeboard.com)

> The_1 domestic_4 cat_1 is_1 a_1 contradiction_5. No_1 other_1 animal_1 has_1 developed_5 such_1 an_1 intimate_5 relationship_1 with_1 humanity_3, while_1 at_1 the_1 same_1 time_1 demanding_1 and_1 getting_1 such_1 independent_2 movement_1 and_1 action_1. The_1 cat_1 manages_1 to_1 remain_1 a_1 tame_8 animal_1 because_1 of_1 the_1 sequence_4 of_1 its_1 upbringing_8. By_1 living_4 both_1 with_1 other_1 cats_1 (its_1 mother_1 and_1 **littermates**) and_1 with_1 humans_1 (the_1 family_1 that_1 has_1 adopted_2 it_1) during_1 its_1 infancy_8 and_1 **kittenhood**, it_1 becomes_1 attached_2 to_1 and_1 considers_1 that_1 it_1 belongs_2 to_1 both_1 species_1. It_1 is_1 like_1 a_1 child_1 that_1 grows_1 up_1 in_1 a_1 foreign_1 country_1 and_1 as_1 a_1 consequence_3 becomes_1 bilingual_8. The_1…

In this case, the cover rate of JACET 8000 is as high as 98%. Although 'littermate' is a difficult word whose meaning we can hardly guess by its morphological constituents, 'kittenhood' would be easily understood as a variation of 'kitten,' which is ranked at the level 7. This shows that JACET 8000 covers almost all of the words in the SAT reading section.

The third sample is taken from the US President's Addressee at the United Nations General Assembly in 2003.

Sample 3. President's Address at the UN Assembly (from whitehouse.gov)

> Last_1 month_1, terrorists_4 brought_1 their_1 war_1 to_1 the_1 United_1 Nations_1 itself_1. The_1 U.N. headquarters_4 in_1 **Baghdad** stood_1 for_1 order_1 and_1 compassion_7 -- and_1 for_1 that_1 reason_1, the_1 terrorists_4 decided_1 it_1 must_1 be_1 destroyed_2. Among_1 the_1 **22** people_1 who_1 were_1 murdered_2 was_1 **Sergio Vieira de Mello**. Over_1 the_1 decades_2, this_1 good_1 and_1 brave_3 man_1 from_1 **Brazil** gave_1 help_1 to_1 the_1 afflicted_8 in_1 **Bangladesh**, **Cyprus**, **Mozambique**, **Lebanon**, **Cambodia**, Central_1 Africa_1, **Kosovo**, and_1 East_2 Timor, and_1 was_1 aiding_1 the_1 people_1 of_1 **Iraq** in_1 their_1 time_1 of_1 need_1. America_1 joins_1 you_1, his_1 colleagues_2, in_1 honoring_2 the_1 memory_1 of_1 **Senor Vieira de Mello**, and_1 the_1…

A mechanical calculation shows that only 82% of the words are covered by JACET 8000. This rate is quite similar to that of the newspapers by Mochizuki (2003b), which was '85%.' However, if we put aside the proper nouns and numerals as in the case of a newspaper article, all words in the Sample 3 are included in JACET 8000.

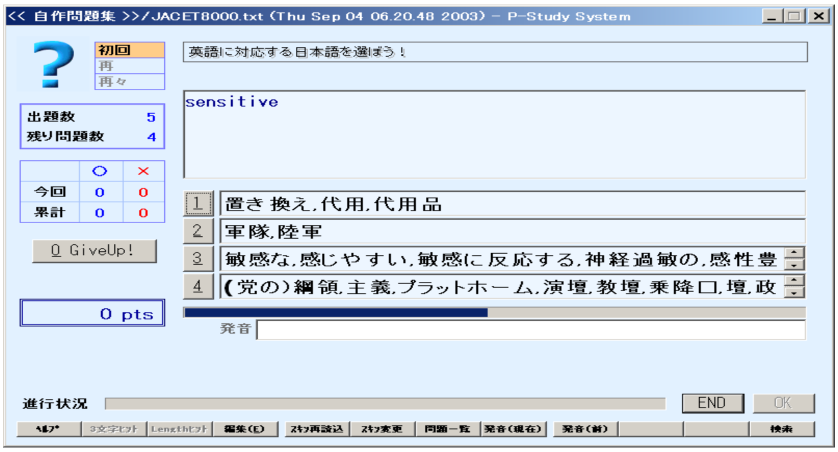
Our analysis of three sample passages proves that JACET 8000 effectively covers the vocabulary in not only easy reading materials but also authentic contemporary English texts.

JACET 8000 IN CLASSROOM USE

As summarized above, JACET 8000 is a word list satisfying a high academic standard, but it is also applicable to classroom activities. Its electronic version is easy to use with other educational PC programs which are free of charge on the Internet (See Ishikawa (2004) for detail). Using

Nishida's freeware, 'Appse,' students in the classroom can easily add word definitions to JACET 8000, as shown in Figure 1. The software creates a kind of interactive file. If learners click an English word in the first column, they are automatically taken to the website of the online dictionary, and can see the detailed word usage explanation and several examples. And if they click 'listen!' button in the second column, they can hear the pronunciation of the word, whose sound date is stored online. And in the third column, they see the definitions or the meanings of the words written in Japanese.

FIGURE 1
JACET 8000 Word Definitions Supplied by Appse



<< 自作問題集 >>/JACET8000.txt (Thu Sep 04 06.20.48 2003) – P-Study System

英語に対応する日本語を選ぼう！

sensitive

1 置き換え, 代用, 代用品
2 軍隊, 陸軍
3 敏感な, 感じやすい, 敏感に反応する, 神経過敏の, 感性豊
4 (党の)綱領, 主義, プラットホーム, 演壇, 教壇, 乗降口, 壇, 政

出題数 5
残り問題数 4

	○	×
今回	0	0
累計	0	0

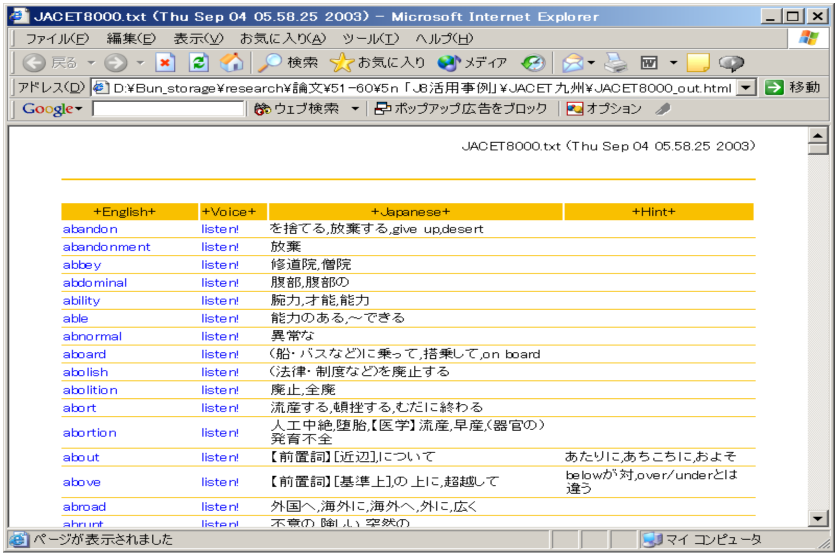
0 pts

With the default setting, the 'Apsse' uses the English-Japanese dictionary data, but if users prepare the machine readable data of dictionary, they can acquire English-English, English- Korean, or English-Chinese outputs.

This software can also create a data file for Takeuchi's 'P-Study System,' shown in Figure 2, a famous software program for vocabulary building. With this CALL program, learners can study the 8000 words in various learning modes such as multiple choice, translation, or typing training from English to Japanese or vice versa.

FIGURE 2

P-Study System with Apsse Data File “Sensitive”


JACET8000.txt (Thu Sep 04 05.58.25 2003) – Microsoft Internet Explorer
ファイル(F) 編集(E) 表示(V) お気に入り(A) ツール(T) ヘルプ(H)
戻る 検索 お気に入り メディア
アドレス(D) D:¥Bun_storage¥research¥論文¥51−60¥5n「J8活用事例」¥JACET九州¥JACET8000_out.html 移動
Google ウェブ検索 ポップアップ広告をブロック オプション
ページが表示されました
マイ コンピュータ

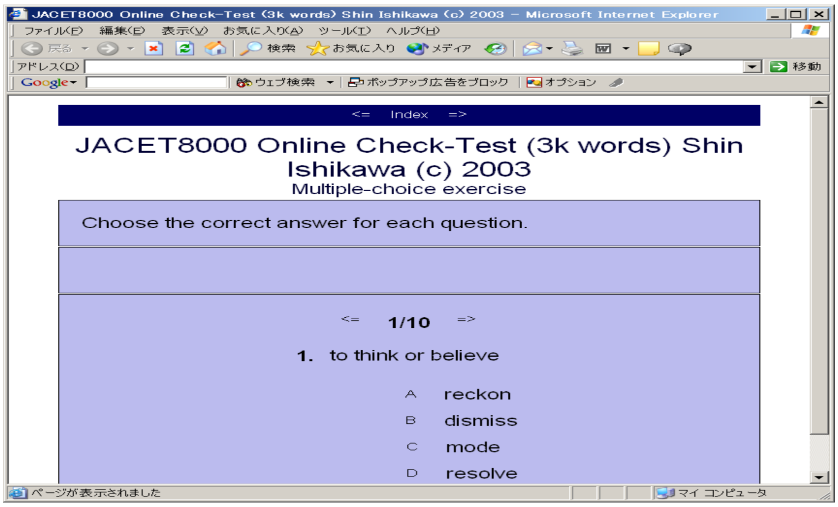
JACET8000.txt (Thu Sep 04 05.58.25 2003)

+English+	+Voice+	+Japanese+	+Hint+
abandon	listen!	を捨てる,放棄する,give up,desert	
abandonment	listen!	放棄	
abbey	listen!	修道院,僧院	
abdominal	listen!	腹部,腹部の	
ability	listen!	腕力,才能,能力	
able	listen!	能力のある,〜できる	
abnormal	listen!	異常な	
aboard	listen!	(船・バスなど)に乗って,搭乗して,on board	
abolish	listen!	(法律・制度など)を廃止する	
abolition	listen!	廃止,全廃	
abort	listen!	流産する,頓挫する,むだに終わる	
abortion	listen!	人工中絶,堕胎,【医学】流産,早産,(器官の)発育不全	
about	listen!	【前置詞】[近辺],について	あたりに,あちこちに,およそ
above	listen!	【前置詞】[基準上],の 上に,超越して	belowが対,over/underとは違う
abroad	listen!	外国へ,海外に,海外へ,外に,広く	
abrupt	listen!	不意の,険しい,突然の	

Appropriateness of vocabulary test is largely influenced by its test designs, and choice of a truly reliable test form is considerably difficult. (See Ishikawa (2003b) for detail.) However, by using ‘Hot Potatoes,’ which creates interactive multiple-choice, short-answer, jumbled-sentence, crossword, matching/ ordering and gap-fill exercises, teachers can easily prepare a kind of check test to see how much the students master the 8000 words, as shown in Figure 3. With ‘Hot Potatoes,’ we can easily make a simple multiple choice test, which serves a wide range of educational purposes.

As briefly introduced above, JACET 8000 can be used to develop English learning computer programs. In combination with various kinds of CALL software, JACET 8000 can drastically change the traditional style of vocabulary learning.

FIGURE 3
Vocabulary Test by JACET 8000 and Hot Potatoes


JACET8000 Online Check-Test (3k words) Shin Ishikawa (c) 2003 - Microsoft Internet Explorer
ファイル(F) 編集(E) 表示(V) お気に入り(A) ツール(T) ヘルプ(H)
戻る 検索 お気に入り メディア
アドレス(D)
移動
Google ウェブ検索 ポップアップ広告をブロック オプション
<= Index =>
ページが表示されました
マイ コンピュータ

JACET8000 Online Check-Test (3k words) Shin Ishikawa (c) 2003
Multiple-choice exercise

Choose the correct answer for each question.

<= 1/10 =>

1. to think or believe

A reckon
B dismiss
C mode
D resolve

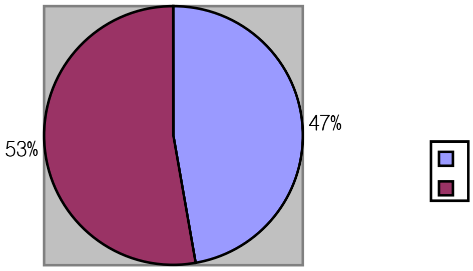
JACET 8000 AND ASIA TEFL VOCABULARY INITIATIVE

JACET 8000 and our experiences of compiling English texts can be shared by many Asian colleagues. It will also be a basis of linguistic comparison of Englishes in Asia. In order to make a sound comparison, we propose three project plans, which are named as Asia TEFL Vocabulary Initiative.

Proposal One: Basing on our local English resources, let’s compile country-based English text corpora, and make their word lists so that we can make a comparison of words of frequent occurrence and patterns of word combination. We can use local English newspapers, English textbooks for Junior & senior high school students, government official English documents, etc. as the sources of domestic English texts. Also, we can supplement or compare our corpus data with much large-scale English corpora such as the British National Corpus, the International Corpus of English - Great Britain, and its Indian, Filipino and Singaporean Components.

In terms of the number of TOEFL examinees, China, Japan and Korea are the three major countries. According to the *TOEFL Test Score and Data Summary, July 2002 to June 2003,* 47% of all TOEFL examinees speak either Chinese, Japanese or Korean as their mother tongues. TOEFL on-line writing test, *ScoreItNow!,* has to do with the second plan.

FIGURE 4
July 2001 to June 2002 TOEFL Examinees



Proposal Two: Let's use the TOEFL on-line writing test, *ScoreItNow!,* to examine the three country students' active knowledge of English vocabulary and use of English.

In students' essays, we often find many errors such as "I lost all my *furnitures*," and "I learn English *since ten years*." Here again, our corpus-based approach is applicable to their compositions. When analyzing these sentences, we can make use of parsing software and concordance programs.

English loanwords in the native language are another index of diffusion of English in Asia. For example, the Japanese sometimes drink *Calpis*, and add *Creap* to their coffee or tea. Toward these Japanese coinages an American professor showed his strong feeling of hatred. To his ear, *Calpis* sounded like "calf piss." *Creap* reminded him of its homophone "creep," and made him itchy.

The third proposal is to research the borrowing from English vocabulary.

Proposal Three: Let's make a comparative study on lexical impact of

English on Asian languages. We can learn much from Görlach et al's "European Anglicism" research project, which draws a map of the general diffusion of English in Europe. Being English professionals in Asia, we can also carry out a similar collaborative research project. As a kick-start, we can have a brainstorming session to gather English loans or variants in our own languages, then examine their pronunciation, spelling and range of meanings.

CONCLUSION

Rapid development of information technology and current trends in economic globalization are so urgent that most Asian people realize a need of a common language by which they can successfully communicate with people of different cultures. Nobody would dispute that English is the language of global communication. However, little attention has been paid to how to compare one variety of English to other varieties of English. In this paper we claim that JACET 8000 and our experiences of compiling English texts can be applicable to comparing varieties of English. Also, basing on our experiences of making a word list, we make three proposals to reveal the nature of our English variety, and to survey English loanwords to examine their influence on our own mother tongue.

THE AUTHORS

Toshihiko Uemura is a professor at Siebold University of Nagasaki, Japan. He got his MA from Aoyama Gakuin University in Tokyo. Having completed his PH. D. coursework, he started teaching at Nagasaki Prefectural Women's Junior College in 1985. Since 1999 he has been teaching English linguistics at Siebold. His recent academic research interests include linguistic phenomena which concern World Englishes and English learning.

Shin'ichiro Ishikawa is an associate professor of English and Applied

Linguistics at Kobe University, Japan. He received his MA from Kobe University and Ph.D. degree from Okayama University. His current research interests include corpus linguistics, lexicography, and literary text analysis.

REFERENCES

Appse. http://www.h5.dion.ne.jp/~kyoun/apsse.html.

Barbera, M. (2003). A dictionary of European Anglicisms: A usage dictionary of Anglicisms in sixteen European languages. *International Journal of Lexicography*, *16*(2), 208-216.

British National Corpus (BNC). http://www.natcorp.ox.ac.uk/.

Committee of Revising JACET Basic Words (Ed.). (2003). *JACET list of 8000 basic words*. Tokyo: Japan Association of College English Teachers.

Committee of Revising JACET Basic Words (Ed.). (2004). *How to make the best of JACET 8000: For educational and research application*. Tokyo: Japan Association of College English Teachers.

Erling, E. (2002). I learn English since ten years: The global English debate and the German university classroom. *English Today*, *70*, *18*(2), 8-13.

Görlach, M. (Ed.). (2001). *A dictionary of European Anglicisms: A usage dictionary of Anglicisms in sixteen European languages*. Oxford: Oxford University Press.

Görlach, M. (2002a). *English in Europe*. Oxford: Oxford University Press.

Görlach, M. (2002b). An annotated bibliography of European Anglicisms. Oxford: Oxford University Press.

Graddol, D. (1997). *The future of English?* London: British Council. http://www.britishcouncil.org/english/pdf/future.pdf.

Hot Potatoes. http://web.uvic.ca/hrd/halfbaked/.

International Corpus of English (Great Britain ICE-GB). http://www.ucl.ac.uk/english-usage/ice-gb/index.htm (Other ICE Components). http://www.ucl.ac.uk/english-usage/ice/avail.htm.

Ishikawa, S. (2003a). Harmonization of a scientific viewpoint and an educational viewpoint in compiling a word list: A case study of 'JACET list of 8000 basic words.' *ASIALEX '03 Tokyo Proceedings*, 72-377.

Ishikawa, S. (2003b). Reliability of various forms of vocabulary test: A pilot study for developing a JACET 8000 vocabulary test. *Papers on Languages and Cultures*, *12*, 67-80. Center for Foreign Languages, Chiba University.

Ishikawa, S. (2004). Teaching vocabulary at colleges with JACET 8000 and various educational software. *How to make the best of JACET 8000*, 7-14.

Laufer, B. (1992). How much lexis is necessary for reading comprehension? In P. Arnaud & H. Béjoint (Eds.), *Vocabulary and applied linguistics* (pp. 126-132). London: Macmillan.

Leech, G., Rayson, P., & Wilson, A. (2001). *Word frequencies in written and spoken English: based on the British National Corpus*. Edinburgh: Pearson Educational Limited. http://www.comp.lancs.ac.uk/ucrel/bncfreq/.

Mochizuki, M. (2003a). JACET 8000 compared with other vocabulary lists. *ASIALEX '03 Tokyo Proceedings*, 378-383.

Mochizuki, M. (2003b). JACET 8000 compared with JACET 4000. *Papers on Languages and Cultures, 12*, 51-57. Center for Foreign Languages, Chiba University.

Murata, M. (2003). Background of revision of JACET word list. *ASIALEX '03 Tokyo Proceedings*, 356-359.

Murata, M., Tono, Y., & Yamada, S. (Eds.). (2003). *ASIALEX '03 Tokyo Proceedings: Dictionaries and language learning – How can dictionaries help human and machine learning?* Tokyo: Asian Association for Lexicography.

Population Division and Statistics Division of the United Nations Secretariat. http://unstats.un.org/unsd/demographic/social/population.htm#pop

P-Study System. http://www.takke.jp/.

Rayson, P., & Garside, R. (2000). Comparing corpora using frequency profiling. *Proceedings of the Workshop on Comparing Corpora, Held in Conjunction with the 38th Annual Meeting of the Association for Computational Linguistics (ACL 2000)* (pp. 1-6).

Shimizu, S. (2003). Data processing for building JACET 8000 vocabulary list. *ASIALEX '03 Tokyo Proceedings*, 360-365.

Sugimori, N. (2003). Using frequency statistics to adjust the rank order of words in JACET list of 8000 basic words. *ASIALEX ’03 Tokyo Proceedings*, 366-371.

Trudgill, P., & Hannah, J. (2002) *4th edition international English: A guide to varieties of standard English*. London: Edward Arnold.

U.S. Census Bureau, International Data Base. http://www.census.gov/cgi-bin/ipc/idbrank.pl

Yun, W., & Jia, F. (2003). Using English in China. *English Today*, *76*, *19*(4), 42-47.

Zhu, H. (2003). Globalization and new ELT challenges in China. *English Today*, *76*, *19*(4), 36-41.