Cognates in Vocabulary Size Testing – a Distorting Influence?

EOIN JORDAN
Xi'an Jiaotong-Liverpool University

Bio Data:
Eoin Jordan (MA, University of Nottingham) currently works as an English tutor at Xi'an Jiaotong-Liverpool University in China. His research interests include vocabulary acquisition, L1 transfer, and English pronunciation.

Abstract
This article examines the issue of cognates in frequency-based vocabulary size testing. Data from a pilot study for a cognate-controlled English vocabulary size test was used to assess whether a group of Japanese university English learners (n=60) were more successful at responding to cognate items than noncognate ones in three 1000 word frequency bands on a Japanese-English translation task. The results showed a statistically significant difference between scores achieved on cognate and noncognate items at the 2000 and 3000 frequency levels, but not at the 1000 frequency level. The findings suggest that cognate items may be easier for test-takers to respond to than noncognate ones of similar frequency, indicating the importance of ensuring that their respective proportions in tests are representative of those inherent in the frequency bands they have been sampled from. It is also argued that such representativeness may best be achieved via a stratified item sampling approach.

Keywords: vocabulary, test, cognates, loan-words, Japanese, English

Introduction

Cognate words, that is those that '[come] naturally from the same root, or [represent] the same original word, with differences due to subsequent separate phonetic development' (OED Online, 2010), are undoubtedly encountered by learners from many L1 backgrounds during their study of English. While it is commonly known that a large number of English words share a root with other European tongues, it is also the case that some more linguistically distant languages have absorbed a great deal of vocabulary from, or that is cognate with, English too. In particular, a body of work by Daulton (1998, 1999, 2003, 2008) suggests that a large proportion of high frequency English vocabulary is cognate for Japanese learners, and asserts that these words can be utilized to assist Japanese students in their language studies. It should be noted that the vocabulary items which Daulton examines are also sometimes described as 'loan words' (as in Kay (1995)). While there is some controversy over

which term is most appropriate, these items are referred to as cognates here in order to maintain consistency with Daulton's work.

The effect that cognateness of a word has on how easy it is to recognize and learn has been investigated in depth by several researchers (De Groot & Keijzer, 2000; Ellis & Beaton, 1993; Hall, 2002; Lotto & De Groot, 1998). Their findings indicate that cognate words are indeed easier to acquire than noncognate ones, which in turn suggests that cognates do have the potential to be utilized effectively in language learning. This viewpoint is endorsed by Nation (2003), who argues that noticing of cognates is a valuable vocabulary expansion strategy.

Despite attention being paid to the effect of cognates on second language learning, there has been relatively little research conducted on the effect that they might have in language testing (although see Meara, Lightbown and Halter (1994) regarding French-English cognates in vocabulary testing). In fact, to the best of the author's knowledge no work has been done at all on this issue in the Japanese context. This is surprising, given that the presence of cognate vocabulary has previously been highlighted as an important problem in some well-known vocabulary tests such as the Vocabulary Levels Test (VLT) (Nation, 1990; Read, 2000) and Yes/No format tests (Eyckmans, 2004). In particular, it is possible that cognates may have a distorting effect on results for frequency-based vocabulary tests where a small number of items are randomly sampled from a large frequency band of words. When the sampling rate is low (for example, 10 items taken from a 1000 word frequency band, as in the Nation's Vocabulary Size Test (VST) (Beglar, 2010)) there is a considerable chance that the proportion of cognate items in the test will not be representative of the proportion of cognates in the frequency band. If cognate items are easier for test-takers than noncognate ones of the same frequency level, then this has the potential to affect results.

This article examines data from an English-to-Japanese translation task that was used as a criterion measure for a cognate-controlled frequency-based English vocabulary size test. A stratified sampling approach to Japanese cognate words was employed in the instrument's construction; thus allowing for a comparison to be made between scores on cognate and noncognate items in several of its frequency levels. It was anticipated that participants would score higher on average for cognate items in each frequency band, which would suggest that these items were indeed easier for respondents.

Method

Participants

The data examined in this article was collected from a total of 61 participants at Ritsumeikan Asia Pacific University in Japan; 59 Japanese undergraduate students and two Japanese English tutors (some Chinese and Korean undergraduate students had also participated in the study, but their results were discarded as it was felt that they may have been influenced by Japanese language difficulties). 56 of the undergraduates were studying on a 'Fundamental English' course, which had a paper-based TOEFL target score of 450, while the remaining 3 were registered on the university's 'Intermediate English' course, which had a target score of 500. The

students participated in the study under the supervision of their tutors during allocated class time, although Intermediate students were also given the option of engaging in an alternative activity during their class, which was the likely cause of the low number of respondents at this level. The two Japanese English tutors had both completed Masters degrees at English language institutions in Britain and America, and they participated in the study in their own time. Participants were not offered any rewards and were required to fill out an online consent form, which was included as part of the research instrument.

Instruments

The instrument used for the study consisted of an online version of a translation task followed by the new vocabulary size test mentioned in the introduction, which had corresponding vocabulary. However, as it is only the data collected from the translation task that is of interest here, a full description of the vocabulary size test is beyond the scope of this article. The translation task consisted of 100 items, sampled to cover the first 5000 words of English, so 20 items represented each 1000 word frequency band. This was expected to be sufficient to measure the vocabulary sizes of Japanese university students, who made up the majority of the participants; previous estimates of Japanese university students' average vocabulary sizes were 2000 (Shillaw, 1995) and 2300 (Barrow Nakanishi & Ishino, 1999). Translation was chosen as a criterion measure in line with Nation (2001), Eyckmans (2004), and Laufer and Goldstein (2004), who all favour it as the most thorough method by which to test receptive vocabulary knowledge.

Word Lists

The translation task was constructed from filtered versions of Paul Nation's (2009) British National Corpus (BNC) word family lists. These are frequency-based lists of word families (Bauer & Nation, 1993), defined in line with Bauer and Nation's Level 6 criteria (Nation, personal communication). Nation's data was filtered using the JACET 8000 lists (Aizawa, Ishikawa & Murata, 2005), which claim to represent the most important 8000 English word families for Japanese students. These are partly based on the BNC, but are also reflective of English teaching materials in Japan. It was intended that this process would remove any 'outliers'; that is, words which were considerably more or less familiar to Japanese learners than Nation's frequency levels indicated.

Each of Nation's frequency banded lists of headwords was inputted into the JACET 8000 Level Marker (Shimizu, 2009). The results, illustrated in Table 1, showed which thousand word JACET 8000 band each of Nation's headwords was located in. A small group of words were found to place very differently on the two lists; one example of this was the word 'confer', which was ranked in the 5000 band of JACET 8000 despite being in Nation's 1k list.

Table 1
Nation's BNC lists divided into JACET 8000 levels

JACET 8000 lists	Nation's BNC lists				
	1k	2k	3k	4k	5k
1000	745	187	15	12	7
2000	162	418	144	52	8
3000	31	172	270	140	60
4000	24	127	148	154	77
5000	8	42	160	142	143
6000	3	23	118	136	137
7000	0	2	41	98	124
8000	0	1	15	74	102
Other	9	26	89	191	342
Unidentified	18	2	0	1	0
Total	1000	1000	1000	1000	1000

The translation task was to consist of 20 headwords from each 1000 word frequency band, which meant that a single item would be representative of 50 words. In line with this, JACET 8000 word bands which had an overlap of less than 50 with any of Nation's individual frequency levels were excluded from that particular level. If included, these words would have been an over-representation of that JACET 8000 band. The exclusions meant that there should have been fewer outliers in terms of word difficulty for Japanese learners of English within each frequency level.

Stratified Sampling of Cognates

In order to ensure that cognates in the word lists were sampled in as accurate proportions as possible, it was first necessary to determine which of the word families featured Japanese cognates as their members and which did not. A list of English cognates for Japanese English learners constructed by Daulton (2003) was used as a starting point for this assessment. Daulton's work was based on one of Paul Nation's previous lists of the most common 3000 word families of English, and he found that a surprisingly high proportion of these appeared to have Japanese loan word equivalents. He highlights that not all of these cognates are equally closely related to their English equivalents, but for the purposes of this study it was felt that the most meaningful division that could be made was between cognates and noncognates; this difference was likely to have more effect on learners' ability to identify a correct form-meaning link than the difference between any two levels of positive cognateness. However the use of Daulton's lists presented two problems: (i)

The lists only covered the initial 3000 word families of English, whereas the translation task in this study was to cover filtered lists of the first 5000 word families, (ii) The first 3000 word families on which Daulton's lists were based were different to Nation's more recent lists.

In response to the two problems highlighted above, it was decided to supplement Daulton's (2003) list with cognate words derived from a Japanese corpus frequency list, and then to ascertain the intersection between this combined list and the filtered lists described in section 2.2.2. . Time and resource constraints meant that it was not possible to carry out an thorough empirical investigation into which Japanese loan words that corresponded with English words on the filtered lists were known by Japanese university students, rather the procedures described here were intended to provide a rough estimate of the number of cognates, with the aim of contributing to the rigorousness of the study.

Assessing which Japanese loan words are widely known among the Japanese population. Just because a Japanese loan word equivalent exists for an English word, it does not follow that this loan word will be in general usage. Likewise, the recent abundant usage of *katakana* (the Japanese script in which most recent foreign loan words from European languages are written) means that even though an internet search may reveal the existence of a few instances of a *katakana* word, this does not mean that it will be known by the majority of the Japanese population. In fact, the Japanese government has recently made efforts to stem the flow of new, difficult to understand loan words into Japanese by proposing alternative phrasings that utilize *kanji* (Chinese characters) to express the same concepts (The National Institute for Japanese Language, 2006). Bearing these factors in mind, one way of identifying loan words that are likely to be well known is to search for them in the upper range of a corpus frequency list; if such words are used frequently then it is likely that they will be familiar to most Japanese speakers.

There are very few publicly available large balanced general Japanese corpora (Goto, 2003; Ueyama, 2006); however a frequency list of lemmas from Serge Sharoff's internet corpus of Japanese (Sharoff, 2009) was considered to be suitable for the purposes of identifying high frequency loan words. As the corpus was assembled through the internet, computer and internet related words such as *kurikku* ('click') and *netowaaku* ('network') were ranked a lot higher than would be expected if it had covered spoken language. One can reasonably expect, however, that the majority of young adults studying English within the education system in Japan are at least to some extent familiar with the internet; thus it seemed plausible to argue that high frequency words on this list were generally well known among the population of young adult English learners.

Loan words in Japanese are usually written in the *katakana* script (Tohsaku, 1993). With this in mind, all of the *katakana* words from the first 10,000 entries of the lemma list were extracted for further analysis. The lemma frequency list actually contained 15,000 entries; however in order to be as confident as possible that the words extracted would definitely be known, only the first 10,000 lemmas were analyzed. Although no clear guidelines could be found on how many lemmas adult native speakers of Japanese are likely to know, the Japanese Language Proficiency

Test (The Japan Foundation, 2008) states that the vocabulary attainment target for the highest level test (Level 1) is knowledge of 10,000 words, and that this level is sufficient to operate in all areas of everyday life. It is not clear whether this refers to individual words, lemmas or word families, but it was felt that extracting *katakana* words from the first 10,000 lemmas of the frequency list would at least provide a conservative estimate of loan words in general usage.

The list of *katakana* words obtained from the corpus frequency list was translated through two online translation sites (Breen, 2010; Google Translate, 2009) and the English translations and *katakana* words checked for consistency. Although a rather subjective measure, any words that had been given English translations which the author felt did not make sense were also checked in the Wisdom Japanese-English Dictionary (Onishi, 2008). This is a widely used dictionary that lists definitions according to frequency of occurrence in an English corpus. Words that were found not to have English cognates were removed at this stage. The list of English translated words was then combined with the list of Daulton's words and put through the Range program (Heatley, Nation & Coxhead, 2002) to check for correspondence with the filtered versions of Nation's 1k to 5k lists on which the translation task would be based. If a word family contained any words that were cognates then that family was classified as cognate, on the grounds that if knowledge about one member could be inferred then this should then allow learners to infer knowledge about the other related members. Table 1 illustrates the numbers of word families designated as cognates.

Calculation of cognate sampling ratios. For reasons relating to the structure of items on the test that the translation task was used as a criterion measure for, sampling ratios of cognates and noncognates for each frequency band were calculated out of 10 rather than out of 20. The resultant ratios (displayed in Table 2) were multiplied by two, and then used to determine the numbers of different item types selected at random from the cognate and noncognate filtered word frequency lists for inclusion in the translation task.

Table 2
Breakdown of Cognates and Noncognates in the Filtered Frequency Lists and TFVST

List	Cognate total	Noncognate total	Translation task cognates (/10)	Translation task noncognates (/10)
Filtered 1k	572	335	6	4
Filtered 2k	466	438	5	5
Filtered 3k	267	573	3	7
Filtered 4k	150	837	2	8
Filtered 5k	64	921	1	9

Presentation
The translation task was divided into four sections of 25 items each in order to match with the format of the vocabulary size test. On each page of the instrument, 25 English words were presented with blank response boxes next to them. Japanese text at the top of the page instructed participants to fill in the most appropriate Japanese translation for each English word into its neighbouring response box. The instructions also stated that they should enter an 'X' in the relevant box if they did not know a translation for a word. After completing each section, participants were required to click on a button at the bottom of the screen that would move them on to the next section. They were not able to advance onto the next section until they had entered something into every response box on their current section. The instructions at the start of the test also informed students that they would not be able to return to completed sections of the test.

Procedures
Students were directed to the instrument via a temporary link using the university's online learning system. The two teachers who helped with administration of the study were briefed on the instrument in advance and gave a short explanation about it to their classes. Students could ask questions if they wanted to and were supervised throughout the whole procedure. The two Japanese English teacher participants were contacted individually. They agreed to participate during their free time and were sent a link to the instrument by email. All participants were told to follow the instructions presented on the screen during the study, which explained what was required of them in full.

Data Analysis
The data was downloaded and translations were then marked in accordance with the lenient marking scheme used by Eyckmans (2004). This meant that levels 1, 2 and 3 of her taxonomy (shown in Figure 1) were accepted as correct.

Figure 1. Marking taxonomy used for the translation task (from Eyckmans, 2004 p.81)

1. Correct translation
2. Correct translation but wrongly spelled or typed
3. Mistakes due to grammatical category
4. Undoubtedly incorrect translation or no response (X)

A lenient marking scheme was felt to be appropriate as the aim was to evaluate whether students had some level of knowledge of the form-meaning link, not to assess grammatical knowledge or Japanese language ability. Translations were allowed if they could be found in either Aizawa et al. (2005) or the Wisdom English-Japanese Dictionary (Inoue & Akano, 2008), or if they were judged to have the same meaning as translations in these sources and were listed in the online ALC database (SPACEALC, 2000). Correct responses were awarded 1 mark, incorrect

responses 0. Items that had not been categorized as cognates during test construction for which a *katakana* loan word was marked as a correct answer were again noted down, then later excluded from comparison analyses between responses for cognate and noncognate items. There were 17 of these items in total, the majority occurring in the lower frequency levels, suggesting that the cognate lists constructed for this test are probably rather conservative estimates of the total number of cognates within the filtered word lists.

Results and Discussion

General Results

Descriptive statistics and the Cronbach Alpha reliability coefficient for the translation task are displayed in Table 3. The Alpha coefficient was sufficiently high, while the mean score and standard deviation suggested that the frequency range of the task was appropriate for the participants. It was also noted that scores on the translation task decreased on average with each of the five frequency bands (see Table 4), suggesting, as expected, that participants were less familiar with lower frequency vocabulary. In this respect the task had performed in a similar manner to other tests of vocabulary size.

Table 3
Descriptive statistics and reliability coefficient for the translation task

N = 60	*M* (/100)	*SD*	Reliability (Cronbach α)
Translation task	45.28	12.95	.94

Table 4
Word frequency level comparisons of scores from the translation task

N=60	***M* (*SD*)**
	Translation task
1000 frequency level (/20)	17.13 (2.40)
2000 frequency level (/20)	9.90 (3.43)
3000 frequency level (/20)	6.73 (3.24)
4000 frequency level (/20)	5.93 (3.41)

5000 frequency level (/20)	5.83 (2.73)

Differences in Performance on Cognate and Noncognate Items

Given the small numbers of cognate items available to be analysed at the 4000 and 5000 frequency levels, comparisons between the performance of cognate and noncognate items were restricted to the first three frequency bands (although it was noted that average scores on the few cognate items in the lower frequency bands were a great deal higher than those for equivalent noncognate items). Table 5 presents a comparison of correct response rates for cognate and noncognate items in each of these frequency levels, which shows that the correct response rate for cognate items was slightly lower at the 1000 frequency level, but considerably higher at the 2000 and 3000 frequency levels.

Table 5

Comparison of correct response rates for cognate and noncognate items at different frequency levels

	***M* (*SD*, *n*)**	
	Cognate items	Noncognate items
1000 frequency level	.84 (.13, 12)	.87 (.14, 6)
2000 frequency level	.73 (.18, 10)	.26 (.21, 10)
3000 frequency level	.76 (.22, 6)	.14 (.19, 10)

To investigate further, a two-way ANOVA was conducted to examine the effects of word frequency level and item type (cognate or noncognate) on correct response rates. Levene's test for equality of error variances suggested that there was homogeneity of variance between groups; thus making the two-way ANOVA viable technique for this analysis. A significant interaction was found between the two effects, F (2, 48) = 15.226, p = .000. Simple main effects analysis showed that there were significant differences between correct response rates on cognate and noncognate items at the 2000 frequency level (p = .000) and 3000 frequency level (p = .000), but not at the 1000 frequency level (p = .737). The lack of a statistically significant difference at the 1000 frequency level may be explained by students' very high average scores in this section, which suggested that most participants were familiar with the vast majority of vocabulary at this level, regardless of whether it was cognate or not. In general, however, the results justified the stratified sampling of cognate items in the test; cognate items appeared to be easier for participants to

answer correctly, and if random sampling had been employed then this may have resulted in unrepresentative proportions of cognates in frequency levels.

The implications of these results are that frequency-based L2 vocabulary size tests for homogeneous L1 contexts, particularly those that have a low sampling rate, run a considerable risk of producing distorted results if they fail to take account of cognate words during item selection. The stratified sampling procedure outlined in this report is one method that can be used to improve the situation by sampling cognates more proportionately; however, this is reliant on the presence of accurate data about cognates, which may not exist for many language pairs. Indeed, the number of valid *katakana* translations provided for supposedly noncognate words in the translation task reported on here suggests that even the lists of Japanese-English cognates used in this study can only be considered as conservative estimates at best. Accordingly, an important focus for future research in this area will need to be the collection of more accurate data on cognates between widely used languages.

Limitations

This report describes a study that was conducted on one small group of Japanese university students and teachers, many of whom had a similar general English level. The effect that a wider and more evenly spread range of English proficiencies among participants would have had on results is not clear; thus the findings here are somewhat limited in terms of their generalizability. A lack of incentive to perform well on the tests may also have resulted in participants not engaging with the material in an earnest manner. Finally, the number of cognate items in the lower frequency bands was, out of necessity, very small, which meant that other properties of these words may have had some influence on results here.

Conclusion

In general mean scores on cognate items were significantly higher than mean scores on noncognate items in the same frequency level. This finding appeared to justify the cognate-controlled design of the test, suggesting that a stratified sampling approach to cognates is likely to produce more accurate estimates of vocabulary size than random sampling of items from frequency bands (particularly when the sampling rate is low). It is the hope of the author that other similar studies are conducted in the future to add further weight to this assertion. The major problem in implementing stratified sampling of cognates in vocabulary size tests for homogenous L1 groups is likely to be a lack of reliable data on which words are, in fact, cognate with L2. The analysis of the data produced in this study suggested that the lists for English and Japanese used here did not cover all of the cognates, particularly at lower frequency levels, and for many other language pairs reliable data may not exist at all. Accordingly, one other area in which further research is needed is the development, using both empirical investigation and corpus data, of reliable lists of cognates for different language pairs.

References

Aizawa, K., Ishikawa, S., & Murata, T. (2005). *JACET 8000 English Basic Words*. Tokyo: Kirihara-shoten.

Barrow, J., Nakanishi, Y., & Ishino, H. (1999). Assessing Japanese college students' vocabulary knowledge with a self-checking familiarity survey. *System*, *27*(2), 223-247.

Bauer, L., & Nation, I.S.P. (1993). Word families. *International Journal of Lexicography, 6*, 253-279.

Beglar, D. (2010). A Rasch-based validation of the vocabulary size test. *Language Testing, 27*, 101-118.

Breen, J.W. (2010). *WWWJDIC: Online Japanese dictionary service*. Retrieved February 17, 2010 from http://www.csse.monash.edu.au/~jwb/cgi-bin/wwwjdic.cgi?9T

Daulton, F.E. (1998). Japanese loanword cognates and the acquisition of English vocabulary. *The Language Teacher, 22*(1), 17-25.

Daulton, F.E. (1999). List of high-frequency baseword vocabulary for Japanese EFL students. *The Internet TESL Journal, 5*(11). Retrieved February 17, 2010 from http://iteslj.org/lists/Daulton-BasewordVocabulary.html

Daulton, F.E. (2003). List of high-frequency baseword vocabulary for Japanese EFL students #2. *The Internet TESL Journal, 9*(3). Retrieved February 17, 2010 from http://iteslj.org/lists/Daulton-BasewordVocabulary.html

Daulton, F.E. (2008). *Japan's built-in lexicon of English-based loanwords*. Clevendon: Multilingual Matters.

DeGroot, A. M. B., & Keijzer, R. (2000). What is hard to learn is easy to forget: the roles of word concreteness cognate status, and word frequency in foreign-language learning and forgetting. *Language Learning*, 50, 1–56.

Ellis, N.C., & Beaton, A. (1993). Psycholinguistic determinants of foreign language vocabulary learning. *Language Learning*, 43, 559-617.

Eyckmans, J. (2004). *Measuring receptive vocabulary size*. LOT (Landelijke Onderzoekschool Taalwetenschap).

Goto, H. (2003). *言語理論と言語資料：コーパスとコーパス以外のデータ*[Linguistic Theories & Linguistic Resources: Corpora and Other Data]. *日本語学* [Japanese Language Studies] *22*, 6-15. Retrieved February 17, 2010 from http://www.sal.tohoku.ac.jp/~gothit/nhnggk0304.html

Hall, C.J. (2002). The automatic cognate form assumption: Evidence for the parasitic model of vocabulary development. *IRAL - International Review of Applied Linguistics in Language Teaching*, *40*(2), 69-87.

Google Translate (2009). *Google translate*. Retrieved February 17, 2010 from http://translate.google.com

Heatley, A., Nation, I.S.P., & Coxhead, A. (2002). *RANGE and FREQUENCY programs*. Retrieved February 17, 2010 from http://www.vuw.ac.nz/lals/staff/Paul_Nation

Inoue, N., & Akano, I. (2008). *The Wisdom English-Japanese dictionary*. Tokyo: Sanseido.

Kay, G. (1995). English loanwords in Japanese. *World Englishes, 14*(1), 67-76.

Laufer, B., & Goldstein, Z. (2004). Testing vocabulary knowledge: Size, strength, and computer adaptiveness. *Language Learning, 54*(3), 399-436.

Lotto, L., & De Groot, A.M.B. (1998). Effects of learning method and word type on acquiring vocabulary in an unfamiliar language. *Language Learning, 48*, 31–69.

Meara, P., Lightbown, P.M., & Halter, R.H. (1994). The effect of cognates on the applicability of YES/NO vocabulary tests. *Canadian Modern Language Review, 50*(2), 296-311.

Nation, I.S.P. (1990). *Teaching and learning vocabulary*. New York: Newbury House.

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

Nation, I.S.P. (2003). The role of the first language in foreign language learning. *Asian EFL Journal, 5*(2), Retrieved February 17, 2010 from http://www.asian-efl-journal.com/june_2003_PN.php

Nation, I.S.P. (2009). *BNC 14k lists*. Retrieved February 17, 2010 from http://www.lextutor.ca/vp/bnc/nation_14/

OED Online (2010). *Oxford English Dictionary*. Oxford University Press. Retrieved January 3, 2011 from http://www.oed.com/view/Entry/35870?redirectedFrom=cognate

Onishi, T. (2008). *The Wisdom Japanese-English dictionary*. Tokyo: Sanseido.

Read, J. (2000). *Assessing vocabulary*. Cambridge: Cambridge University Press.

Sharoff, S. (2009). *Large corpora used in CTS*. Retrieved February 17, 2010 from http://corpus.leeds.ac.uk/list.html

Shillaw, J. (1995). Using a word list as a focus for vocabulary learning. *The Language Teacher, 19*(2), 58-59.

Shimizu, S. (2009). *JACET 8000 level marker*. Retrieved December 15, 2009 from http://www01.tcp-ip.or.jp/~shin/J8LevelMarker/j8lm.cgi

SPACEALC (2000). *Eitarou on the web*. Retrieved February 17, 2010 from http://www.alc.co.jp/

The Japan Foundation (2008*). The Japanese-language proficiency test*. Retrieved February 17, 2010 from http://www.jlpt.jp/e/

Tohsaku, Y. (1993). *Yookoso!: Invitation to contemporary Japanese*. New York: McGraw-Hill.

Ueyama, M. (2006). Evaluation of Japanese Web-based Reference Corpora: Effects of Seed Selection and Time Interval. In Baroni, M., & Bernardini, S. (Eds.), *Wacky! Working papers on the web as corpus*. Bologna: GEDIT. Retrieved February 17, 2010 from http://wackybook.sslmit.unibo.it/wackycontents.html

Appendix

List of words used in the translation task

SECTION A
ADVANCE, CHARACTER, DRY, EASY, FARM, GERMANY, GROW, HAND, HERE, INDUSTRY, INFORM, MAJOR, NEW, PARTY, PROCESS, PROMOTE, PURE, SIT, SITE, SOCIAL, SUGGEST, THEN, TIME, WINE, WISE.

SECTION B
BIBLE, CASUALTY, CHAPEL, CONTRIBUTE, CRITERION, CUSHION, DIMENSION, FETCH, FORMAL, HILL, INTELLIGENCE, JEANS, LIBERAL, MANUAL, NEVERTHELESS, ORCHESTRA, PUBLISH, REPLY, RIVER, SKY, SPIN, TREMENDOUS, TRIVIAL, ULTIMATE, VEGETABLE.

SECTION C
ANTIQUE, ARTIFICIAL, BAIL, BEE, BUBBLE, CEASE, DELIBERATE, EMPIRE, FIN, GALLON, IRONY, ISRAEL, IVY, MUTUAL, OUTRAGE, PALM, PUNISH, RAVE, RECITE, REVEAL, SHATTER, SOAP, TRACTOR, VERIFY, WHEELCHAIR.

SECTION D
ADMINISTER, BLEND, BLOUSE, BROOM, DEADLINE, DEFICIENCY, DIVINE, EMIGRATE, FLUID, FOG, GASP, INTERCEPT, JAZZ, JURY, MINIATURE, MULTITUDE, PHYSICS, PIERCE, PUBLICIZE, RENDER, RETREAT, SNAKE, TESTAMENT, TRADESMAN, UNDERWEAR.