Creating a Corpus-Based Daily Life Vocabulary for TEYL

Kiyomi CHUJO
College of Industrial Technology, Nihon University, Japan.

Kathryn OGHIGIAN
Center for English Language Education for Science and Engineering, Waseda University.

Masao UTIYAMA
National Institute of Information and Communications Technology, Japan.

Chikako NISHIGAKI
Faculty of Education at Chiba University in Japan.

Bio Data:
Kiyomi Chujo is an associate professor at the College of Industrial Technology, Nihon University, Japan. She completed her Ph.D. on vocabulary selection for English education at Chiba University, Japan in 1991. Her current research interests are vocabulary learning and the pedagogical applications of corpus linguistics such as data-driven learning.

Kathryn Oghigian teaches technical writing and academic reading at the Center for English Language Education for Science and Engineering, Waseda University. She earned her MA degree from the University of British Columbia in Modern Language Education in 1997, and her interests are focused on corpus-based applications to language acquisition.

Masao Utiyama is a senior researcher of the National Institute of Information and Communications Technology, Japan. His main research field is natural language processing. He completed his doctoral dissertation at the University of Tsukuba in 1997. His current research interests are exploring models of natural languages and their practical applications.

Chikako Nishigaki is a professor in the Faculty of Education at Chiba University in Japan. She completed her doctoral dissertation at Chiba University in 1992. She teaches applied linguistics to undergraduate and graduate students and has published extensively on EFL vocabulary and listening.

Abstract

The purpose of this study has been to create a list of children's everyday vocabulary in English which will provide a foundation for daily life vocabulary for Japanese elementary school students and which will complement and augment existing English vocabulary currently taught in Japanese junior and senior high schools. Vocabulary words were taken from the CHILDES spoken corpus and picture dictionaries, and were ranked statistically with an outstanding-ness score based on a log likelihood keyword analysis and a selection probability score based on an adapted form of range. It was found that the identified words are at the appropriate grade level (grades 1 to 3), that the semantic content areas are grade-appropriate and complement the semantic categories of junior and senior high school (JSH) vocabulary, and that this vocabulary supplements JSH vocabulary in text coverage over 18 activities.

Keywords: daily life vocabulary, TEYL, picture dictionary, corpus, CHILDES

Introduction

In Japan, an initiative began in 2002 to teach English to young learners (TEYL) and when a new course of study is fully implemented in 2011, English language activities will become compulsory for fifth and sixth graders. The Ministry of Education, Culture, Sports, Science and Technology (MEXT) wrote the overall objective of English activities "to form the foundation of pupils' communication abilities" (MEXT, 2009) through "conducting conversational activities wherein students can be exposed to daily expressions and terms in English" (Butler & Takeuchi, 2008, p. 69). In anticipation of this reform, MEXT produced a textbook in 2008 called *Eigo No-to* or *English Note* for the fifth and sixth grade curricula. A recent vocabulary analysis of *Eigo No-to* showed that it contained an estimated 386 words, and that 8.1% of these were higher than the U.S. 8th grade level (Chujo & Nishigaki, 2010). While *Eigo No-to* is no doubt a useful resource for teachers, its word selection raises interesting questions about defining 'daily life vocabulary,' the optimal number of words for a curriculum, the most appropriate target for grade level, and how this vocabulary would complement or overlap with the vocabulary currently taught at the junior and senior high school levels. In addition, *Eigo No-to* is not a mandated textbook, and educators are expected to develop their own syllabuses and supplemental materials. With or without this resource, most primary school teachers generally have neither the experience nor the skills necessary for teaching English and they need effective

teaching material that will be successful and motivating so that these early language-learning experiences not only support TEYL but also will become a basis for learning at the secondary level and beyond. This study addresses this need by creating a 1,000-word corpus-based vocabulary of daily life words in English. In this paper, *daily life vocabulary* is defined as the words relevant to the everyday experience of children and young language learners and is used interchangeably with *everyday words* or *everyday vocabulary.*

Literature Review

The Need for Daily Life Vocabulary

Theoretical and empirical research in EFL in Japan suggests that teaching daily life words to elementary-aged children can be highly beneficial for EFL learners (Ito, 2000; Kuno, 1999; Saku & Honda, 2004; Shirahata, 2004) and teaching these kinds of words also meets with the Japanese government's TEYL guidelines (MEXT, 2009) which state that English relevant to children's everyday life should be taught in public elementary schools. Many researchers in Japan have emphasized that this vocabulary is considered to be the core vocabulary of college students and college graduates (Hamano, 1989; Horiuchi, 1976; Inoue, 1985), and the lack of this vocabulary is often felt by teachers and students who go to English-speaking countries for a short stay to experience daily life in native speakers' homes (Inaoka et al.*,* 1988; Tsuruta, 1991). Chujo, Hasegawa, and Takefuta (1994) documented this vocabulary gap in a study comparing the vocabulary coverage of Japanese and American textbooks over eighteen specific language activities. They compiled a 14,694-word list generated from American basal K through 8th grade readers called the *Ginn Reading Program* (Clymer, Venezky, & Indrisano, 1982), and a 3,483-word list generated from the textbooks most widely used in Japanese secondary schools from the 7th through the 12th grades. They found that the American textbook vocabulary covered all activities evenly, but the Japanese textbook vocabulary showed a gap in daily life vocabulary coverage, focusing instead on, for example, student conversations, travel phrases and TOEFL vocabulary. In another study, Hasegawa and Chujo (2004) investigated a series of three Japanese secondary school textbooks used in each of the past three

decades and found that while there have been slight improvements in daily life vocabulary coverage in each ten-year revision of the same textbook series, there was still a lack of daily life words necessary for survival in English. Other researchers have also pointed out that these words in particular are not sufficiently covered in Japanese English textbooks taught in junior and senior high schools (hereafter, JSH) (Inoue, 1985; Mouri, 2004). For example, students rarely learn words such as *drawer, refrigerator, trash,* and *glue* from English textbooks used in JSH (Nishigaki, Chujo, & Oghigian, 2009). Finally, Jin'nai (2003) reported that educators in secondary schools are expecting TEYL to provide the vocabulary currently not taught in Japanese secondary schools.

Daily Life Vocabulary Sources

A century ago when Jespersen, in his *How to Teach a Foreign Language* (1904), stated that the beginner has use for only the most everyday words, "the problem that faced the textbook writer who wished to follow Jespersen's precepts was how to know [exactly] which were 'the most everyday words'" (as cited in Hornby, 1967, p. 41). While it is possible to identify high frequency words in English from general and specialized corpora, to date there have been no known studies done to create this type of children's vocabulary using statistical extraction tools such as log likelihood (Dunning, 1993). It should be noted from the outset that as a general corpus, the British National Corpus (BNC) has been shown to be inappropriate for using unchanged as the basis for syllabus design for EFL or ESL learners in primary or secondary schools because "[t]he BNC is predominantly a corpus of British, adult, formal, informative language, and most English learners in primary and secondary school systems are not British, are children, and need both formal and informal language for both social and informative purposes" (Nation, 2004, pp. 3-4). Ishikawa (2005, p. 44) demonstrated that in the BNC, the rank of words familiar to Japanese schoolchildren such as *notebook, eraser, blackboard, pocket* and *chime* is low and stated that the high frequency words derived from the BNC are weak in identifying young children's familiar everyday vocabulary.

There are currently very few children's corpora available (Danielsson & Mahlberg,

2003, p.4). Because the Japanese TEYL curriculum addresses conversational activities, our focus is on spoken children's corpora, and to date we have identified three: the Polytechnic of Wales Corpus (1978-1984) of children's speech in play sessions and interviews; the Moe, Hopkins, and Rush (1982) corpus of spontaneous conversations with first grade children; and the CHILDES (Child Language Data Exchange System) corpus of conversations with young children (2000). In 2006, Chujo, Utiyama, Nishigaki, Nakamura, and Yamazaki used the *log likelihood* statistic to extract and examine the outstanding vocabulary of these three spoken corpora. This statistical process is known as a *keyword analysis.* Scott (1997, pp. 236-243) defined a *keyword* as a "word which occurs with unusual frequency in a given text" and proposed a method of identifying keywords in text by using the chi-square and log likelihood statistics. Chujo et al. obtained keywords that are used in children's corpora statistically more frequently than in general English by comparing each word's frequency in the children's corpora to its frequency in the BNC general-usage adult spoken list of 9,477 words. They found that the CHILDES corpus contained basic verbs, colorful nouns and adjectives relevant to a young child's everyday world. The PoW corpus contained words limited to the play sessions, games, and interview topics, and the Moe et al. corpus contained words related only to limited subjects. Based on the findings of this 2006 Chujo et al. study, the CHILDES has therefore been identified as an appropriate source for this current study.

In addition to a children's corpus, researchers in Japan agree that picture dictionaries are a vital resource of everyday words (Inoue, 1985; Kittaka, 2000; Matsumura, 2004; Shiina, Chujo, & Takefuta, 1988), and Nishigaki, Chujo, and Iwadate (2005) confirmed that they contain a high level of everyday words. Some picture dictionaries define their specific goals for featuring this vocabulary; for example, *The Basic Oxford Picture Dictionary* (Gramer, 2003) targets "language that is essential for the development of the beginning learner's survival skills" and *The Sesame Street Dictionary* (Hayward, 2004) provides "words that appear frequently in beginning reading books and in a young child's everyday world." Thus, in addition to the CHILDES corpus, we have targeted and included data from picture dictionaries as a

source of everyday words.

Research Questions

Previous studies indicate that there is a need to construct a daily life word list relevant to the everyday experience of children and young language learners for TEYL education at the elementary level in Japan, and that there is also a need for filling in the gap of daily life vocabulary not taught in Japanese secondary schools. Although no single corpus exists to provide a comprehensive selection of this type of vocabulary, the CHILDES corpus and picture dictionaries have been identified as appropriate sources of TEYL vocabulary. The purpose of this study is to create a daily life word list from the CHILDES corpus and picture dictionaries and examine it to determine if it is appropriate for TEYL. Specifically:

1. Are the identified words at the appropriate TEYL grade level?

2. What semantic categories are represented, and how are these distributed over various types of daily life activities?

3. How does this daily life word list compare to existing JSH vocabulary, i.e., does it improve text coverage of everyday words as a supplement to JSH textbooks?

Method

Source Lists

CHILDES. From the CHILDES (Child Language Data Exchange System) spoken data, ten sets of American English native speaker children's speech data ranging from age 2 to age 10-11 (grade 5) were chosen and downloaded[1]. The 129,326 different words in this 1.29 million-word corpus were lemmatized to extract all base forms using the CLAWS7 tag set (1996), that is, inflectional forms such as *cat-cats* and *go-goes-went-gone-going* were listed under the base word forms of *cat* and *go*. All proper nouns and numerals were identified by their POS (part of speech) tag and deleted manually because statistical measures mechanically identify these words as technical words (Scott, 1999). Next, to create a pedagogically applicable list, all unusual or infrequent words (i.e., those occurring only once) were excluded. This process yielded a 4,161-word list.

In accordance with the Chujo et al. 2006 study discussed earlier using the log

likelihood statistic, the 4,161-word CHILDES list was compared with the BNC general-usage adult spoken list of 9,477 words to statistically identify which words are *outstandingly* used in children's speech, compared to that of adults. A score for '*outstanding-ness*' was assigned to indicate the level of use by children compared to that of adults. This procedure provided an outstanding-ness score for each of the 4,161 CHILDES words, and we ranked those words in ascending order according to the 'outstanding-ness' score.

Picture Dictionaries. Twenty picture dictionaries for both native speaking children and ESL/EFL learners published by major overseas publishers in the U.S., England, Australia, Singapore and Hong Kong, and ten picture dictionaries published in Japan were collected. They are listed in Appendix A. The selection criteria for these picture dictionaries were as follows: (a) they contain more than 500 entries; (b) authors state that they provide everyday words; (c) authors state that they have pedagogic value; and (d) they are available.

The words contained in each of the thirty picture dictionaries were manually typed or scanned optically and then reformatted into thirty individual lists. In addition, each list was identified as having been published in Japan and or overseas, thus there were ten lists for ten Japan-based dictionaries and twenty lists for twenty non-Japan-based dictionaries. Next, each word list was lemmatized, and proper nouns and numerals were excluded from each list manually. The number of different words in the twenty dictionaries published abroad totaled 4,691 (Picture Dictionary List 1) and that of ten Japan-based dictionaries was 3,897 (Picture Dictionary List 2), yielding a combined total of 5,259 words (Picture Dictionary List 3).

In picture dictionaries, each individual word is presented with a picture, usually without a context or sentence. An analysis of picture dictionary data therefore would not (and did not) produce a normal frequency list as would be obtained from an analysis of text data. Because of this, the criteria of 'frequency of occurrence' often used in studies was not applicable. In addition, there are no stated criteria for each author's inclusion or ranking for each entry word. Since it is likely these were decided intuitively based on expertise (no explicit rationale was given for any dictionary), we

used 'range' to express a numerical consensus. For example, words that appeared in all twenty overseas picture dictionaries were referred to as 'range 20.' If the size of all the picture dictionaries were the same, we could say the words having a wider range are more important than those with a smaller range. However, there was a difference in size between the picture dictionaries. For example, *Word by Word* (Molinsky & Bliss, 1995) contains 2,554 different words and *Ladybird Picture Dictionary* (Taylor, 2004) contains only 608 different words. In this case, it is reasonable to assume that a word found in a smaller picture dictionary is more important than one found in a larger picture dictionary.

In order to generalize the idea of the 'range,' we proposed an adapted form of range called 'selection probability,' which enables the importance of a word to be weighted in favor of a word that is found in a smaller sized picture dictionary (see Chujo et al., 2005). For that purpose, we assigned a probability to each word[2]. This resulted in a selection probability score for each word on each of the three lists (Picture Dictionary Lists 1, 2 and 3[3]) and each list was ranked in ascending order according to the probability score.

Creating a Ranked Daily Life TEYL Vocabulary Master List

We next integrated the four lists (the CHILDES list and the three picture dictionary lists) into one "Daily Life TEYL Vocabulary List" with each word showing one ranked score. To do this, we used the following procedure:

1. It was important to handle the picture dictionaries published outside Japan and in Japan separately when we calculated the selection probability. In the exploratory phase of our research, we examined the selection probability scores of the picture dictionary lists and found that the words from picture dictionaries published within and outside of Japan were based on different cultural views. For example, Japan-based picture dictionaries included words such as *curry, persimmon, leapfrog* and *squid* as everyday vocabulary which would be useful in a Japanese context, but not necessarily outside of Japan, i.e., for students or teachers living abroad, or for a wider Asian EFL audience. Therefore, we used the selection probability scores of the picture dictionaries published in the U.S., England, Australia, Singapore, and Hong Kong, so

that these daily life words would rank higher than the Japan-based daily life words. Thus, while Japan-specific words such as *persimmon* and *leapfrog* would be included in the master list, these would be ranked much lower than words encountered in situations abroad such as *asleep* or *dollar*. Therefore, we have two basic statistics assigned to each word in one master list ("Daily Life TEYL Vocabulary List"): the selection probability score for the picture dictionaries and the CHILDES log likelihood 'outstanding-ness' scores.

2. Next, we calculated an average of the rankings of the selection probability scores for the picture dictionaries and the rankings of the CHILDES log likelihood 'outstanding-ness' scores, and then ranked the words in ascending order. This list is available on the web at http: //www5d.biglobe.ne.jp/~chujo/.

Evaluating the Word Lists

In order to determine the pedagogical appropriateness for TEYL, the words on the Daily Life TEYL Vocabulary List (hereafter 'TEYL List') were evaluated with regard to grade level, semantic content and distribution, and JSH text coverage. These procedures are discussed below.

1. Determining the grade level of the TEYL vocabulary. In order to understand at what U.S. grade level these words would be understood by native English speaking (American NS) children, the list of 5,259 TEYL words was compared to *The Living Word Vocabulary* (Dale & O'Rourke, 1981) and the *Basic Elementary Reading Vocabularies* (Harris & Jacobson, 1972). *The Living Word Vocabulary* includes more than 44,000 items and each presents a percentage score for those words or terms familiar to students in grade levels 4, 6, 8, 10, 12, 13, and 16. (Note that grades 13 through 16 denote four years at the college or university level.) The *Basic Elementary Reading Vocabularies*, with 7,613 different words appearing in a selection of textbooks widely used in 1970 in grades one through six of the elementary school, was used for determining the (U.S.) grade levels of reading vocabulary for the first, second, and third grade levels. Using these control lists, we calculated the average grade level for ten different list sizes from the top-500 to the top-5,000 TEYL words. Although we acknowledge that these sources are dated, we were able to determine

grade levels for all the words appearing on our list, since generally these basic words have not changed over time, for example, *pencil, chair, book,* and *toy*. In addition, there is no contemporary comparable resource that we are aware of[4].

2. Determining the semantic categories of the TEYL vocabulary. Tom McArthur's *Longman Lexicon of Contemporary English* (1981) classifies over 15,000 entries under a set of fourteen semantic fields such as *life and living things*, and *people and the family*. In this study, we used these 15,000 entries in the fourteen semantic fields to make it possible to cluster words in a word list into groups of different semantic fields[5]. Some polysemous words, for example *nail,* belong to two semantic fields: *the body*; and *substances, materials, objects, and equipment*. Therefore the total number of semantic fields is larger than the number of words.

To confirm that the TEYL list includes grade-appropriate concepts such as animals, food, school, nature, and the home environment, we compared the distribution of the semantic fields of the first 500 words from the TEYL list to the fourteen semantic fields of words in the JSH textbook vocabulary. Although most of the first 500 TEYL words do not appear in the JSH textbook vocabulary, there was overlap. In order to examine distribution, first those words that appear both in the TEYL list and the JSH vocabulary were deleted from the first 500 TEYL list. In order to maintain 500 words, this TEYL list was supplemented with words from the second 500 TEYL list so that there were a total of 500 TEYL words, and this modified "Top 500 TEYL (Ver. 2) list" was then compared to the fourteen categories.

3. Determining the JSH text coverage of the TEYL vocabulary. Finally, to understand how the TEYL vocabulary compares to existing JSH vocabulary, text coverage was calculated. A JSH vocabulary list, containing 3,950 different base words, was compiled from the 41,112-word top selling series of textbooks, the *New Horizon 1, 2, 3* series (Tokyo Shoseki, 2002) and the *Unicorn I, II & Reading* series (Bun'eido, 2003) currently used in Japanese secondary education. We wanted to see how well this JSH vocabulary covered various activities, and how this compared to the coverage provided by the TEYL vocabulary. For this purpose, five 1,500-word text samples of eighteen language activities were used from a previous study (see Chujo et

al., 1994)[6]. These activities include nine text categories used in spoken language such as daily conversation, survival conversation, movies, medical conversation with nurses and doctors, economic news, business talk, a radio program, and TOEFL listening sections; and nine text categories used in written language such as a cooking article, an everyday word dictionary[7], a woman's magazine, science news, a business letter, a computer manual, a science book, a novel, and a *Time* magazine. The sources are listed in Appendix B.

Text coverage was calculated by counting the number of the words known in the text, multiplying this number by 100 and then dividing by the total number of words in the text. Using the formula p = (the number of words covered in the activity text by the TEYL list words)/ (total number of words in the activity text) x 100, we calculated the targeted vocabulary coverage percentage learners might reasonably be expected to obtain along with the acquisition of the JSH level vocabulary and TEYL vocabulary.

Results and Discussion

Research Question 1: Evaluating Grade Level

The results of a comparison of the TEYL words with *The Living Word Vocabulary* (Dale & O'Rourke, 1981) and the *Basic Elementary Reading Vocabularies* (Harris & Jacobson, 1972) are shown in Table 1. In addition to the average grade level, we calculated the standard deviation (SD) of each of the top-500 to the top-5,000 TEYL words to measure how far any number (grade level score) is from the middle. For example, a SD of 2.0 allows that the grade level may range from the average grade level ±2.0.

Table 1

Vocabulary Size and Average Grade Level

Vocabulary size	Average grade level	SD
500	2.4	1.2
1,000	2.9	1.6
1,500	3.6	2.4
2,000	4.2	3.1
2,500	4.9	3.8
3,000	5.3	4.1
3,500	5.6	4.3
4,000	5.9	4.6
4,500	6.2	4.7
5,000	6.0	4.6

We can see a clear tendency for a steady increase in grade level with the change of vocabulary size, and an increase in the SD, which means the grade levels are less stable among each vocabulary strata as the vocabulary size increases toward 5,000 words. We can see that the first 500 words and the first 1,000 words are generally understood by third grade students, with a SD of 1.2 and 1.6, respectively. The levels increase systematically: The 2,000 word strata are generally known by fourth grade students, the 4,000 word strata by fifth graders and the 5,000 word strata by sixth grade students. We also see that a larger vocabulary has a larger SD compared to a smaller vocabulary. Thus we can expect to obtain a more reliable grade level when the vocabulary size is smaller.

It is notable that the average grade level of the first 500 and 1,000 words remains stable at 2.4 and 2.9 respectively, and that they have a smaller SD (less than 2.0) compared to the larger vocabulary strata. This procedure allowed us to identify an optimal number of words for a smaller working word list. Japanese educators

(Takefuta & Suikou, 2005; Ono, 2005) advocate allotting 500 words or 500 to 1,000 words to TEYL in primary education based on the estimation that the required size an adult EFL learner's vocabulary for practical communication activities is 7,000 to 8,000 words (Takefuta & Suikou, 2005, p. 60). Therefore, we limited the TEYL list to 1,000 words. In terms of practical application, we can say that these first 500 words and/or the first 1,000 words might be the most appropriate and useful vocabulary size for selecting daily life words for beginner level TEYL students and that they are within the elementary school range, that is, grades 1 through 3. Therefore as a more pedagogically useful vocabulary list, we have 500 or 1,000 grade-appropriate TEYL words from the original list of 5,259 words.

We can confirm that the log likelihood and selection probability statistics we used to rank the words were reasonable with regard to grade appropriateness. And from Appendix C, we can clearly see that appropriate words for the lower grades are listed in the first 500 words.

Research Question 2: Evaluating Semantic Content and Distribution

By comparing the TEYL word list to the *Longman Lexicon of Contemporary English* (McArthur, 1981) we were able to determine that it included words in each of the fourteen semantic categories. Figure 1 represents the distribution of semantic fields for 500 TEYL words (Ver. 2) and the JSH vocabulary. The percentage of TEYL words classified into each semantic field is shown with black bars, and the percentage of JSH words is show with gray bars.

We can see the top semantic fields of the TEYL words are: (a) *life and living things*; (b) *substance*, *materials*, *objects*, *and equipment*; (c) *buildings*, *houses*, *the home*, *clothes*, *belongings*, *and personal care*; (d) *entertainment*, *sports*, *and games*; (e) *movement*, *location*, *travel*, and *transport*; and (f) *food*, *drink*, and *farming*. We can say that the TEYL words (for example, *shoe, cat, car,* and *chair*) generally relate to concrete concepts belonging to semantic fields appropriate to the developmental level of the students.
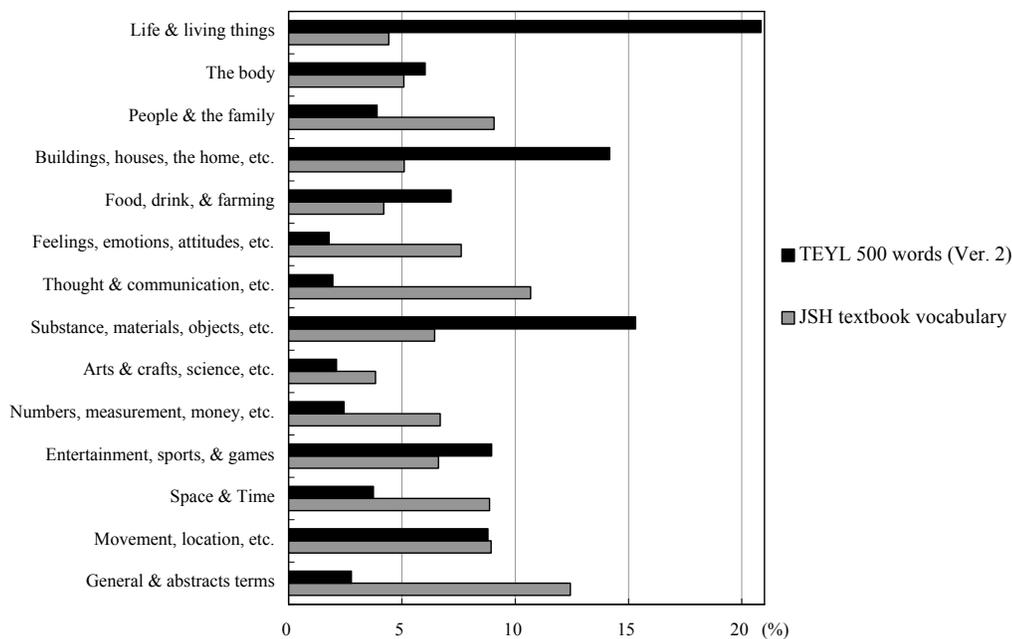
Figure 1. A comparison of percentages of the 500 TEYL words and JSH textbook vocabulary by semantic field.

On the other hand, the top semantic fields of the JSH textbook vocabulary are: (a) *general and abstracts terms*; (b) *thought and communication, language, and grammar*; (c) *people and the family*; (d) *space and time*; (e) *movement, location, travel, and transport*; and (f) *feelings, emotions, attitudes, and sensations*. JSH students "are able to think beyond the immediate context in more abstract terms" (Pinter, 2006, p. 7), and this is reflected in the semantic categories. Overall, from this observation we can see the TEYL words can provide elementary level students with grade appropriate concepts relevant to a child's everyday world.

Research Question 3: Evaluating Text Coverage

Finally, to understand how the TEYL vocabulary compares to existing JSH vocabulary, text coverage was calculated and the results are shown in Figure 2. The percentage of text coverage for the JSH textbook vocabulary over each activity is shown by gray bars; and the JSH textbook vocabulary supplemented by the modified 500 TEYL words (Ver. 2) is shown by black bars. Looking at the graph, we can see the ineffectiveness of the JSH textbook vocabulary, mainly because of its limited scope. Since the JSH texts are for grades 7 through 12, it's appropriate that the

coverage is rather low for adult language activities such as medical conversations with doctors, or reading science news and *Time* magazine. However, the most notable point is that there is a lack of important daily life words in the JSH texts. We can see that the addition of the 500 TEYL words resulted in the improvement of text coverage for 'Everyday words' from 53.3% to 70%. The TEYL is an important supplement, although there would be benefit from further improvements.
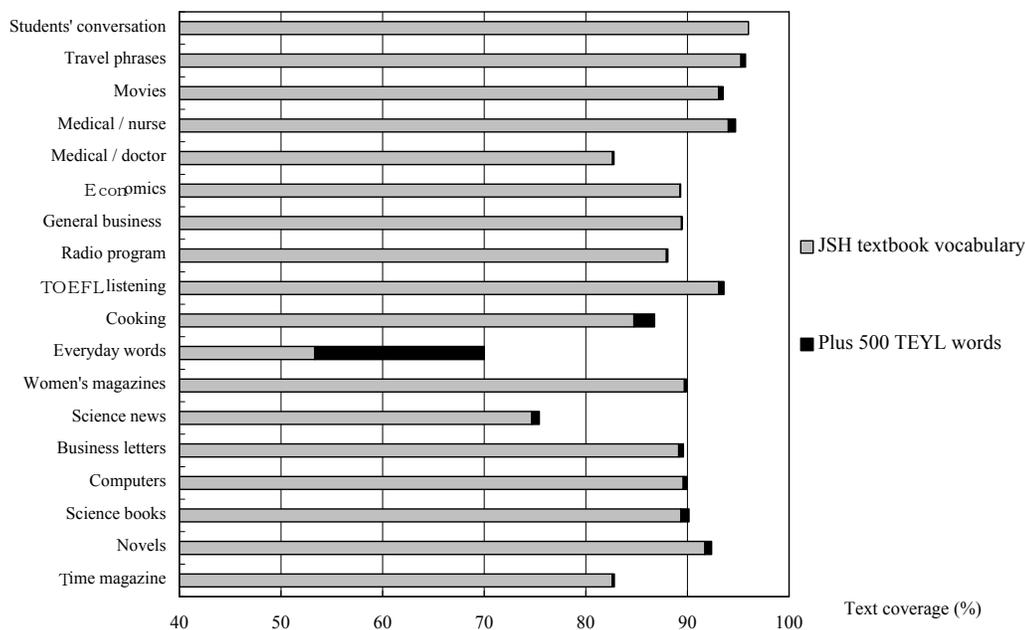
Figure 2. Text coverage of Japanese textbook vocabulary with/without 500 TEYL words over 18 activities texts.

Conclusion

From a review of the literature, we understand that there is a need to construct a word list for TEYL education at the primary level in Japan, and that there are no known studies which have done so. Not only is this type of everyday vocabulary essential to young learners as a basis of language knowledge, it is essential for filling in the gap of vocabulary not taught in Japanese junior and senior high schools, and Japanese secondary level educators expect a word list that will address this lack. In addition, this is important vocabulary for Japanese or other Asian students who travel to English-speaking countries.

In this study, 1,000 words were statistically selected from a children's spoken

corpus and from picture dictionaries and were found to be appropriate with regard to grade level, semantic content and text coverage. Although other TEYL lists might be generated from other sources other than CHILDES and picture dictionaries, we hope this list contributes to the body of work in TEYL language teaching, and that it or lists similar to it will be considered when elementary and JSH textbooks are revised by MEXT over the coming years. Additionally, the methodology used to generate the TEYL list may be of interest to readers outside of Japan. To determine if the TEYL list is useful in other contexts, educators can calculate text coverage calculation by replacing the JSH textbook vocabulary list with the vocabulary from another [Asian] textbook. This list is accessible online at http: //www5d.biglobe.ne.jp/~chujo/eng/index.html. E-learning software programs and gaming devices in four languages (Chinese, Korean, English, Japanese) based on this TEYL list are under development for a broader Asian EFL audience, and an *Ara Karuta* card set (Nishigaki et al., 2009) is currently available.

References

Bun'eido. (2003). *Unicorn I, II & Reading* (Series). Tokyo: Bun'eido.

Butler, Y. G. & Takeuchi, A. (2008). Variables that influence elementary school students' English performance in Japan. *The Journal of Asia TEFL*, *5*(1), 65-95. Retrieved from http://www.asiatefl.org/journal/main15.php

CHILDES (Child Language Data Exchange System) (2000). See MacWhinney, B. (2000).

Chujo, K., Hasegawa, S., & Takefuta, Y. (1994). Nichibei eigo kyoukasho no hikaku kenkyuu kara [A comparative study on English textbook vocabulary in Japan and the U.S.]. *Gendai Eigo Kyouiku, 29* (12), 14-16.

Chujo, K., Nishigaki, C., Utiyama, M., Iwadate, H., & Yamazaki, A. (2005). Eigo ejisho no goi [The vocabulary of selected picture dictionaries]. *Journal of the College of Industrial Technology*, *Nihon University*, *38*, 77-104.

Chujo, K., Utiyama, M., Nishigaki, C., Nakamura, T., & Yamazaki, A. (2006).

Kodomo hanashi corpus no tokuchougo chuushutsu ni kansuru kenkyuu [Extracting outstanding children's spoken words from children's spoken corpora]. *Journal of the College of Industrial Technology*, *Nihon University*, *39*, 65-78.

Chujo, K. & Nishigaki, C. (2010). Shougakkou *Eigo No-to* no goi bunseki [A vocabulary analysis of the *Eigo No-to* resource book for elementary schools], *English Corpus Studies, 17*, 115-126.

CLAWS7 (1996) [Computer software]. Lancaster: Lancaster University. Retrieved from http://ucrel.lancs.ac.uk/claws7tags.html

Clymer, T., Venezky, R. L., &. Indrisano, R. (1982). *The Ginn reading program*. Lexington, Massachusetts: Ginn and Company.

Dale, E., & O'Rourke, J. (1981). *The living word vocabulary*. Chicago: World Book-Childcraft International, Inc.

Danielsson, P. & Mahlberg, M. (2003). There is more to knowing a language than knowing its words: Using parallel corpora in the bilingual classroom. *English for Specific Purposes World, 3*(6), 1-11. Retrieved from http://www.esp-world.info/contents.htm

Dunning, T. E. (1993). Accurate methods for the statistics of surprise and coincidence. *Computational Linguistics*, *19* (1), 61-74.

Hamano, M. (1989). Motto nichijou seikatsu goi wo [More everyday words!]. *Gendai Eigo Kyouiku, 26* (9), 6-7.

Harris, A. J., & Jacobson, M. D. (1972). *Basic elementary reading vocabularies*. New York: The Macmillan Company.

Hasegawa, S., & Chujo, K. (2004). Gakushuu shidou youryou no kaitei ni tomonau gakkou eigo kyoukasho no jidaiteki henka [Vocabulary size and efficacy within three serial JSH English textbook vocabularies created in accordance with revised Course of Study guidelines]. *Language Education & Technology*, *41*, 141-155.

Hiebert, E. H., & Kamil, M. L. (2005). *Teaching and learning vocabulary*. Mahwah, NJ: Lawrence Erlbaum Associates, Publishers.

Horiuchi, K. (1976). Teiji junjo to shiyou hindo [The order of presentation and frequency of occurrence]. *Gendai Eigo Kyouiku, 13* (6), 4-5.

Hornby, A. S. (1967). Vocabulary control: History and principles. In W. R. Lee (Ed.), *E.L.T. selections: Articles from the journal English language teaching 1* (pp. 40-45). London: Oxford University Press.

Inaoka, N., Kato, Y., Kubono, M., Kumai, N., Tsuji, H., Nakamura, Y., & Hasegawa, K. (1988). Chuukou renkei wo toraeta jishu kyouzai no kaihatsu: Tsukukoma picture dictionary no sakusei (1) [Developing student-initiated teaching material to relate junior high school education to senior high school education: The development of the Tsukukoma picture dictionary (1)]. *Tsukuba Daigaku Fuzoku Chuugaku Koutougakkou Kenkyuu Houkoku, 27,* 131-146.

Inoue, N. (1985). Eibei youji muke kyouiku tosho no goi chousa [A survey on words used in British/American educational books for small children]. *Teikoku Gakuen Kiyou, 11,* 109-126.

Ishikawa, S. (2005). Frequency and familiarity in compiling the English word list for children. *IEICE Technical Report, TL 25,* 43-48.

Ito, K. (2000). *Shougakkou eigo: Redii. Gou* [Elementary school education. Ready? Go!]. Tokyo: Gyousei.

Jespersen, O. (1904). *How to teach a foreign language*. London: Allen & Unwin.

Jin'nai, Y. (2003, August). *Shougakkou eigo katsudou wo ikasu shou-chuu renkei no arikata* [How to relate elementary school education to junior high school education to make the best use of English activities at elementary school]. Paper presented at The 29th Japan Society of English Language Education Conference, Sendai: Miyagikyouiku University.

Kittaka, S. (2000). Goi gakushuu kyouzai toshiteno picture dictionary no yuuyousei: Keiyoushi no kanten kara [The usefulness of a picture dictionary as vocabulary teaching material: From the view of adjectives]. *Journal of Kisarazu National College of Technology, 33,* 107-114.

Kuno, Y. (1999). *Konna fuu ni hajimete mitewa? Shougakkou eigo* [How about starting elementary school English like this?]. Tokyo: Sanseido.

MacWhinney, B. (2000). *The CHILDES Project: Tools for analyzing talk. 3rd Edition. 2: The Database.* Mahwah. NJ: Lawrence Erlbaum Associates. Retrieved from http://childes.psy.cmu.edu/data/

Matsumura, T. (2004). Eigo seikatsu goi shidou no jissai [The practical side of teaching English everyday words]. *The 30th Japan Society of English Language Education Conference Proceedings* (pp. 262-265). Nagano: Shinshuu University.

McArthur, T. (1981). *Longman lexicon of contemporary English*. Essex: Longman.

MEXT (2009). Japanese Ministry of Education, Culture, Sports, Science and Technology. The Course of Study Foreign Languages. Retrieved from http://www.mext.go.jp/component/a_menu/education/micro_detail/__icsFiles/afiel dfile/2009/04/21/1261037_12.pdf

Moe, A., Hopkins, C., & Rush, T. (1982). *Vocabulary of first-grade children*. Springfield: Charles C. Thomas Publisher.

Mouri, K. (2004). *Eigo no goi shidou: Anote konote* [Teaching English vocabulary: This way and that]. Hiroshima: Keisuisha.

Nation, P. (2004). A study of the most frequent word families in the British National Corpus. In P. Bogaards & B. Laufer (Eds.), *Vocabulary in a second language* (pp. 3-13). Amsterdam: John Benjamins Publishing Company.

Nishigaki, C., Chujo, K., & Iwadate, H. (2005). Ejisho de manabu nichijou-seikatsu goi [Daily-life vocabulary learnable from picture dictionaries]. *The 26th Japan Association for the Study of Teaching English to Children Conference Proceedings* (pp. 21-24). Kasugai City: Chubu University.

Nishigaki, C., Chujo, K., & Oghigian, K. (2009). *Ara Karuta* [*Oh! Karuta*]. Tokyo: Kairyudo.

Ono, H. (2005). Shougakkou ni okeru mini tsuku eigo gakushuuhou no kaihatsu [Developing effective English learning methods at primary schools]. *The 26th*

Japan Association for the Study of Teaching English to Children Conference Proceedings (pp. 66-69). Kasugai City: Chubu University.

Pinter, A. (2006). *Teaching young language learners*. Oxford: Oxford University Press.

PoW (The Polytechnic of Wales Corpus) (1999). The ICAME Corpus Collection on CD-ROM, version 2, 1999 [Data file]. Retrieved from http://icame.uib.no/newcd.htm

Saku, M., & Honda, K. (2004). Shougakkou eigo kyouiku ni okeru goi kyouzai no kaihatsu ni mukete [Toward developing vocabulary teaching material for English education at elementary school]. *The 30ᵗʰ Japan Society of English Language Education Conference Proceedings* (pp. 598-599). Nagano: Shinshuu University.

Scott, M. (1997). PC analysis of key words and key key words. *System*, *25*(2), 233-245. doi:10.1016/S0346-251X(97)00011-0

Scott, M., (1999). Wordsmith tools [Computer software]. Oxford: Oxford University Press. Retrieved from http://www.lexically.net/wordsmith/

Shiina, K., Chujo, K., & Takefuta, Y. (1988). Youji jidou muke etangoshuu no bunsekiteki kousatsu [The analytical observations of English picture dictionaries for young children]. *JASTEC Journal, 7*, 17-27.

Shirahata, T. (2004, July). *Nihon no kodomo ni totteno eigo kyouiku towa* [English education for children in Japan]. Paper presented at the 42ⁿᵈ Japan Association for Language Education and Technology Conference, Hakata, Japan.

Takefuta, Y., & Suikou, M. (2005). *Kore kara no daigaku eigo kyouiku* [College English education in the future]. Tokyo: Iwanami Shoten.

Tokyo Shoseki. (2002). *New Horizon 1, 2, 3* (Series). Tokyo: Tokyo Shoseki.

Tsuruta, Y. (1991). Sunde shitta seikatsu goi no iryoku [It was not until I experienced living abroad that I learned the power of daily life vocabulary]. *Eigo Kyouiku*, *39*(13), 46-49.

Zeno, S., Ivens, S., Millard, R., & Dubburi, R. (1995). *The educator's word frequency*

guide. Brewster, NY: Touchstone Applied Science Associates, Inc.

Footnotes

1. From the "English-American Corpora" section of CHILDES, ten sub-corpora titled Bliss, Bohannon, Brown, Carterette & Jones, Evans, Garvey, Gathercole, Kuczaj, Tardif, and Van Kleeck, were chosen based on the subjects' age range and data collection situation. For details on these corpora, please consult the 'English-American Corpora' section (http: //childes.psy.cmu.edu/data/) as well as a general introduction to the CHILDES (http: //childes.psy.cmu.edu/).

2. The selection probability of a word extracted from 20 dictionaries is defined as follows. To select a word from a dictionary, we first select a dictionary, di (i=1…20), from the 20 picture dictionaries. Thus, the selection probability of di, $P(di)$, is 1/20. Next, we select word w from di. Suppose that di has $W(di)$ words, the selection probability of w given di, $P(w|di)$, is $1/W(di)$. Thus, the selection probability of di and w, $P(w,di)$, is $P(di)*P(w|di) = (1/20)*(1/W(di))$. Note that $P(w,di)$ is 0 if w is not included in di. We add the selection probability of di and w, $P(w,di)$, for the 20 dictionaries to calculate $P(w) = P(w,d1) + P(w,d2) + ... + P(w,d20)$. The selection probability, $P(w)$, is a generalization of range. Suppose that all the dictionaries are the same size, i.e., $W(d1) = W(d2) =... = W(d20) = K$, where K is a constant. Then, if the range of word w is r, then $P(w) = r * (1/20)* (1/K) = r *$ constant. Thus, $P(w)$ is proportional to r. The selection probability weights words in smaller dictionaries more heavily than words in larger dictionaries. For example, if $W(d1)$ =1000 and $W(d2) =$ 2000, then $P(w,d1) = (1/20)*(1/1000)$ and $P(w,d2) = (1/20)*(1/2000)$. Thus, $P(w,d1) > P(w,d2)$. This is because a word contained in a smaller dictionary is more important than a word contained in a larger dictionary.

3. It was noted that all of the CHILDES words were already included in the Picture Dictionary List 3.

4. The rationale for using *The Living Word Vocabulary* (LWV) is explained by Hiebert (2005, pp. 252-253):

> … the time frame within which it was validated make[s] the LWV a less-than-ideal resource for use with students in the early part of the 21st century. At the present time, however, the LWV is the only comprehensive, existing database on students' familiarity with word meanings…[and furthermore]…"Because of the shortcomings in the LWV system, an additional resource [is necessary] …for decisions of inclusion or exclusion on grade-level lists….

Because the LWV assigned grade level 4 to grade level words from grades 1-4, we used an additional resource to evaluate the grade levels of those words, and allotted each word to grade 1, 2, 3, and 4. Although the newer Zeno et al. (1995) was available, we wanted to use a resource from a similar time frame as the LWV, and therefore chose Harris & Jacobson (1972).

5. Although the UCREL Semantic Analysis System (http://ucrel.lancs.ac.uk/usas/) is effective for analyzing text, we learned that the precision of the semantic tagging for a word list might not be as precise as that for a text (personal communication with P. Rayson, January 2, 2007).

6. In order to ensure the reliability of the results and to confirm the results were

not dependent on the type of text, it was necessary to replicate the previous study (Chujo et al., 1994). Thus we used the same eighteen sets of vocabulary for the 18 language activities used in the 1994 study, even though the materials may be somewhat dated.

7. It should be noted that because a picture dictionary was included in this control list, it was not included as one of the thirty picture dictionaries chosen for the study.

APPENDIX A

Selected Picture Dictionaries

Title	Author	Publisher	Year	Words
The Longman Picture Dictionary American English	Ashworth, J. & Clark, J.	Pearson Education Ltd. , Harlow, Essex	1993	1,066
Word by Word	Molinsky, J. S. & Bliss, B.	Prentice Hall Regents, Englewood Cliffs	1995	2,554
Just Look'n Learn ENGLISH Picture Dictionary	Hochstatter, D.	Passport Books, Lincolnwood, IL	1996	1,274
Scholastic First Dictionary	Levey, S. J.	Scholastic Inc., New York	1998	1,614
The Oxford Picture Dictionary for Kids	Keyes, R. J.	Oxford University Press, New York	1998	761
Smile Picture Dictionary	Barraclough, C.	Macmillan Heinemann, Oxford	1999	748
The Oxford Picture Dictionary for the Content Areas	Kauffman, D. & Apple, G.	Oxford University Press, New York	2000	1,207
Word by Word Primary phonics picture dictionary	Molinsky, J. S. & Bliss, B.	Pearson Education Ltd., White Plains, NY	2000	863
First Word Study Dictionary	Turton, N.	Learners Publishing Pte Ltd., Godown, Singapore	2001	905
My Big Word Book	Priddy, R. et al.	Priddy Bicknell, New York	2002	876
My World A First Picture Encyclopedia	Picthall, C. & Gunzi, C.	Reader's Digest Children's Publishing Inc., Pleasantville, NY	2002	681
Oxford First Dictionary	Goldsmith, E.	Oxford University Press	2002	1,362
Disney My First 1000 words	Feldman, T.	Disney Press, New York	2003	1,193
Disney Picture Dictionary	Feldman, T. & Benjamin, A.	Disney Press, New York	2003	971
First Picture Dictionary	Oliver, A.	Hinkler Books, Dingley, Victoria, Australia	2003	766
Longman Children's Picture	Graham, C.	Longman Asia ELT, Quarry Bay,	2003	739

Dictionary		Hong Kong		
Longman Photo Dictionary of American English	Summers, D. et al.	Longman, Harlow, Essex	2003	2,195
The Basic Oxford Picture Dictionary (2nd ed.)	Gramer, M.	Oxford University Press, New York	2003	1,114
Picture Dictionary	Taylor, G.	Ladybird Books Ltd., London	2004	608
The Sesame Street Dictionary	Hayward, L.	Random House, New York	2004	1,174
WORD BOOK: E-de Mite Oboeru Eitango	Kuno, Y.	Borgnan, Tokyo	1993	1,004
ABCD Book: Hajimete Deau Eigo-no Jiten	Yoneyama, E.	Sekai Bunka Publishing Inc., Tokyo	1995	1,031
English-Japanese Picture Dictionary	Toda, K. & Herring, A. K.	Toda Design Kenkyuushitsu, Tokyo	1999	1,126
Sanseido Word Book 1	Hatori，H., Kuno, Y. & Kaizaki, Y.	Sanseido Publishing Co., Ltd., Tokyo	1999	828
ALC Picture Dictionary: 2000 Words for Kids	Kuno, Y.	ALK Co., Ltd., Tokyo	2000	1,704
Sanseido Word Book 2	Kuno, Y. & Arthur, B.	Sanseido Publishing Co., Ltd., Tokyo	2000	1,142
Kodomo Eigo Jiten	Tsuruta, K.	Kodansha Ltd., Tokyo	2001	813
English for You	Yasuyoshi, I.	Seibido Shuppan Co., Ltd., Tokyo	2001	691
NOVA Illustrated English Dictionary	NOVA	Nova Corporation, Tokyo	2004	2,701
Hajimete Eitango Jiten	Gakken	Gakken Co., Ltd., Tokyo	2004	920

APPENDIX B

Eighteen Language Activities and Their Sources

Spoken Language			
Activities	Sources	Types	Tokens
Students' Conversation	Svartvik, J. & Quirk, R. (1980). *A corpus of English conversation*, Lund: C. W. K. Gleerup, 804-811.	292	1555
Travel Phrases	Church, N. & Moss, A. (1983). *How to survive in the U.S.A.: English for travelers and newcomers*. Cambridge: Cambridge University Press, 125-132.	358	1586
Movies	Tucker, Broadcast News, My Fair Lady, The Unbearable Lightness of Being, Mother Teresa.	429	1570
Medical / Nurse	David, A. & Crosfield, T. (1976). *English for nurses*, Burnt Mill, Harlow: Longman 5, 10-11, 93-94, 97, 106-107, 111-112, 114-116.	326	1557
Medical / Doctor	Cassidy, L. & Wagatsuma, T. (1982). *English conversation for physicians*, Tokyo: Medical View Co., LTD., 66-74, 79.	521	1520
Economics	ABC News and Gloview Co., Ltd. (1982). *Abc special Dan Cordtz's Econocast*.Tokyo: ABC News and Gloview Co., Ltd., 80, 82, 84, 88, 90, 92, 94.	437	1519
General Business	Howe, B. (1988). *Visitron: The language of meeting and negotiations*. Burnt Mill, Harlow: Longman, 52-54, 62-64, 72-73.	403	1531
Radio program	Goken. (1987). *FEN news file volume 7*, Tokyo: Goken, 8-60.	409	1529
TOEFL Listening	Sharpe, P. J. (1989). *Barron's how to prepare for the TOEFL*, Sixth Edition. New York: Barron's Educational Series, Inc., 559-563.	484	1534
Written Language			
Cooking	*Ladies' Home Journal, 104*(5), 1988, 66, 120, 128, 142-145.	566	1527
Everyday Words	Parnwell, E. C. (1988). *The new Oxford picture dictionary*. New York: Oxford University Press, 2-50.	1091	1500
Women's Magazines	*Savvy*, 1988, March, 22-23, 26.	475	1525
Science News	*Science News, 136*(26), December 23 & 30, 1989, 416-422.	538	1512

Business Letters	*Webster's guide to business correspondence*, Springfield: Merriam-Webster Inc., 1988, 265- 289.	477	1509
Computers	Naiman, Arthur (1988-89). *The Macintosh bible*, Second Edition. Berkeley, California: Goldstein & Blair, 31-37.	678	1497
Science Books	Hawking, S. (1988). *A brief history of time: from the big bang to black hole.* New York: Bantam Books, 1-6.	446	1509
Novels	Christie, A. (1968). *N or M?* New York: Dell Publishing Co., Inc. , 5-10.	361	1536
Time	*Time*, No.14, April 2, 1990, 11-12.	494	1506

APPENDIX C

The 1,000 Daily Life TEYL Vocabulary (in Order of Rank)
Note that number beside each word indicates the grade level according to Dale &
O'Rourke (1981) and Harris & Jacobson (1972). Any words not appearing in either
resource are denoted by '*'

Word	#	Word	#	Word	#	Word	#	Word	#	Word	#	Word	#
baby	1	tree	1	brown	1	bath	3	rubber	3	neck	2	sled	2
pencil	3	write	2	dinosaur	4	broom	3	mirror	3	friend	1	whistle	2
fish	1	crayon	4	catch	1	pull	2	kiss	3	alphabet	4	flute	4
milk	2	mouse	2	chicken	2	jacket	3	around	1	show	1	apartment	2
chair	2	green	4	clock	2	motorcycle	4	chin	3	spaghetti	4	chalk	4
dog	4	cup	2	do	1	wet	1	piece	2	wait	4	camp	3
finger	2	fall	1	down	4	bone	2	man	1	pet	1	bake	2
mouth	2	hammer	3	beach	3	make	1	bubble	3	rice	3	tricycle	4
toy	1	bread	3	refrigerator	3	peach	4	rhinoceros	6	grandmother	2	bracelet	4
car	1	banana	3	floor	2	tire	2	hide	2	noise	2	bucket	4
nose	2	dress	1	glue	4	kitchen	2	name	1	soap	4	castle	3
ball	4	can	1	wheel	2	shovel	2	oven	2	shadow	3	grapefruit	4
cheese	4	soup	2	pink	2	wash	2	pepper	4	right	1	mix	3
orange	3	read	1	little	1	food	1	ribbon	3	helicopter	3	needle	3
hot	2	telephone	2	draw	3	bat	4	octopus	4	circus	2	leopard	4
shoe	1	bee	1	throw	2	thumb	4	star	2	toast	4	wind	2
book	1	cream	2	sugar	3	puppet	4	strawberry	4	sunglass	6	uncle	2
apple	2	kite	2	drum	3	chocolate	4	bicycle	3	lobster	4	toothbrush	4
hat	1	birthday	1	open	2	train	1	on	1	panda	4	dig	2
paper	2	sing	1	rope	2	turn	2	bite	3	drawer	3	kid	4
table	2	rabbit	1	zebra	4	owl	2	cousin	3	dollar	2	gray	2
eat	1	girl	1	look	1	worm	4	off	1	purse	4	goldfish	4
head	1	sock	4	moon	3	tomato	4	bottle	2	noodle	4	pizza	4
snake	4	house	1	up	1	swing	3	no	1	outside	2	mail	2
eye	2	balloon	1	store	1	baseball	3	giant	2	bow	3	wake	2
lion	2	puppy	2	camel	4	yogurt	10	belt	3	fix	2	eagle	3
hand	1	bear	1	I	1	flag	3	sink	4	snail	4	beetle	4
hair	1	big	4	arm	2	honey	2	pumpkin	3	aunt	2	polar	8
cat	1	knife	3	blanket	3	hippopotamus	4	penguin	4	tent	2	nail	3
juice	4	fork	3	butter	2	lettuce	4	wear	2	mom	4	earring	4
bird	1	swim	2	fire	1	fruit	1	chick	4	shark	4	teacher	4
truck	1	brother	2	break	2	hungry	2	out	1	knock	2	screwdriver	4
bed	1	go	1	tie	2	feather	2	slide	2	recorder	4	closet	4
tooth	2	drink	2	light	1	stove	3	bathroom	4	cave	3	pour	3
elephant	2	horse	1	pen	2	pajama	4	room	2	violin	4	stand	2
ride	1	cook	2	fly	1	knee	3	dragon	2	stick	2	skate	2
jump	1	whale	3	tall	2	gorilla	4	home	1	kick	3	trash	4
boat	1	sit	1	nut	4	tongue	3	horn	2	mailbox	4	watermelon	4
flower	2	push	2	door	2	alligator	4			beard	3	want	1
doll	2	lunch	2	cry	1	black	1	roll	2	eyebrow	4	dump	4
frog	3	sheep	2	cut	1	suitcase	4	spider	3	like	1	pin	3
water	1	giraffe	3	bean	3	hop	1	goose	2	crocodile	4	flamingo	6
red	1	plate	3	brush	3	bridge	2	pocket	1	sad	2	notebook	6
ice	1	carrot	3	clown	2	rug	3	street	1	wood	2	father	1
blue	4	mother	1	bowl	2	ant	4	basketball	4	donkey	3	heavy	2
butterfly	3	tractor	2	over	1	basket	2	elbow	4	guitar	4	rocket	1
monkey	2	leg	1	towel	4	clean	2	salad	4	cherry	4	face	2
boot	2	comb	3	circle	3	clothes	4	pear	4	mushroom	4	penny	1
animal	1	puzzle	3	rain	1	dish	2	fun	4	bell	2	king	2
color	1	kangaroo	4	goat	1	mask	4	mountain	2	medicine	3	pineapple	4
sandwich	3	pillow	3	zoo	1	seal	4	dinner	2	bike	1	cage	1
tiger	2	cookie	2	hole	2	seed	2	wing	2	desk	3	bulldozer	4
boy	1	box	1	string	2	meat	3	scare	2	hug	3	raccoon	2
shirt	2	tail	2	barn	1	ladder	2	piano	3	hurt	2	kind	1
cake	1	school	1	dance	2	pea	4	next	1	help	4	thirsty	4
egg	2	grape	4	candle	2	walk	1	coffee	3	napkin	4	mustache	4
watch	2	block	2	button	2	close	2	monster	4	letter	1	hill	1
airplane	1	snow	2	fast	1	picnic	1	last	1	hold	1	farmer	2
yellow	1	pant	6	sweater	4	put	1	caterpillar	4	lamb	2	jelly	4
a	1	cereal	4	corn	2	sky	3	hit	2	song	2	busy	2
cow	1	glass	2	glove	2	drop	1	flashlight	4	happy	1	shape	2
spoon	4	rock	2	squirrel	2	shell	3	drive	2	pick	2	turkey	3
duck	1	sister	1	bag	1	farm	1	snowman	2	sneaker	4	garage	2
scissor	4	dirty	2	candy	3	popcorn	4	saucer	4	peanut	1	sick	3
play	1	purple	3	white	1	you	1	hard	1	garbage	4	here	1
toe	3	cold	4	coat	1	dad	2	grasshopper	4	magic	2	plane	4
see	1	picture	1	climb	2	blow	2	pie	2	parrot	3	laugh	1
sleep	1	kitten	1	hamburger	4	dry	2	slipper	4	sandbox	4	touch	3
foot	1	turtle	4	mitten	4	top	2	good	1	cheek	3		
paint	1	pig	1	people	2	grass	1	take	1	xylophone	6		
ear	2	breakfast	2	sand	2	pan	1	window	1	eraser	4		
game	1	umbrella	4	sun	1	movie	4	deer	2	rooster	3		

step	1	plant	2	ocean	3	tight	3	lock	3	chase	2	chess	4
sharp	3	wolf	2	cloud	3	shelf	3	museum	3	painting	4	straw	3
sailboat	4	cone	3	screw	4	ink	4	miss	1	crash	3	alone	2
playground	4	cube	6	tummy	4	ready	1	rattle	4	chop	3	log	3
pot	3	arrow	3	forget	2	sea	2	sword	4	paste	4	bunny	2
spell	4	yo-yo	4	lamp	3	yard	2	smile	2	quilt	4	suntan	4
jewelry	4	sweep	3	reindeer	4	raincoat	4	skip	4	gum	4	he	1
how	1	air	2	back	1	reach	2	shoot	3	stethoscope	10	barber	4
tell	1	dentist	4	elevator	2	toaster	4	feed	2	baker	4	ballet	4
fence	2	wrist	4	camera	3	seashell	4	doughnut	4	x-ray	4	tugboat	4
cactus	4	stair	4	wheat	4	peel	4	thread	3	diaper	4	east	3
seesaw	4	ceiling	3	statue	3	teapot	4	apron	3	bacon	4	north	3
taste	3	wipe	3	pretend	3	whisper	2	orangutan	13	zip	4	south	3
fold	3	loud	2	dirt	3	lightning	3	fingernail	4	gymnastic	6	pliers	4
pail	2	sidewalk	2	dark	1	hurry	1	empty	2	hockey	4	tow	6
backpack	*	park	2	icicle	4	starfish	4	crescent	8	sharpener	4	track	2
wheelbarrow	4	lip	4	hay	3	dune	8	flipper	4	fountain	3	squeeze	4
ship	2	pony	1	bump	2	sew	3	cotton	4	curl	3	quart	4
koala	1	grow	2	ketchup	4	snack	4	ape	4	warm	2	crack	3
bug	3	wastebasket	4	summer	2	desert	4	scarecrow	2	bakery	4	windmill	4
rake	4	surprise	1	costume	4	dessert	4	thermometer	4	dustpan	4	officer	4
jar	2	stir	4	scarf	4	robot	4	mixer	4	compass	4	curly	4
real	2	stop	1	crawl	2	subway	6	today	2	melon	4	taco	*
wagon	1	fur	3	waterfall	4	taxi	4	asleep	3	parade	2	yesterday	3
leaf	4	tower	2	bring	1	milkshake	4	raspberry	4	hang	2	jigsaw	4
lizard	4	river	2	trunk	2	wrap	3	supermarket	4	ranch	2	secret	3
toothpaste	4	spaceship	4	ski	4	cucumber	6	beautiful	2	bend	3	backyard	3
rainbow	4	helmet	4	it	1	bark	1	rattlesnake	4	hawk	4	stool	3
celery	4	ax	4	shake	2	roller	4	burn	4	stem	3	tortoise	4
tennis	4	pole	3	paw	2	upside	3	shampoo	4	armchair	4	toolbox	*
fat	1	iron	2	roof	2	witch	3	insect	3	web	4	mash	4
chimney	3	bus	1	grandfather	2	powder	4	belong	2	scream	3	dandelion	4
quiet	2	melt	3	treasure	3	clam	6	skeleton	4	butcher	4	mop	4
dime	3	tool	3	another	1	bull	4	poor	2	cape	3	cushion	4
carpenter	4	cracker	4	hike	4	yell	2	unicorn	12	dot	2	roast	3
careful	2	envelope	4	fry	3	where	1	vegetable	3	diamond	3	soon	1
clap	3	jungle	3	bounce	2	beaver	3	blouse	4	shave	4	knight	4
ladybug	4	pop	2	cool	3	telescope	3	undershirt	4	driver	3	ping-pong	4
knot	4	cap	3	sneeze	3	puddle	4	waffle	4	hear	1	tulip	4
necklace	4	pipe	3	couch	4	bench	3	fireman	4	sleeve	4	piggy	4
what	1	spill	3	nest	2	tunnel	3	hook	3	drill	3	almost	2
rectangle	6	lemon	4	mud	4	scratch	3	astronaut	4	snowflake	4	volcano	4
finish	2	acrobat	4	chimpanzee	4	checker	4	cartoon	4	marshmallow	4	sausage	4
come	1	rat	4	key	3	story	1	under	1	waiter	4	tiny	2
canoe	4	jet	3	soda	4	coin	3	stay	1	jellyfish	4	win	2
race	1	heel	3	bib	4	skunk	3	aquarium	4	grain	3	wizard	4
triangle	4	grandma	4	hen	1	sofa	4	daddy	4	swan	3	wallet	4
doctor	2	pretty	2	silly	2	minute	2	yawn	4	trumpet	4	hippo	4
supper	2	some	1	wrench	4	blueberry	2	daisy	4	who	1	hydrant	6
count	2	wall	2	need	2	sponge	4	inside	2	bookcase	4	soy	12
long	1	soccer	10	fit	2	crib	4	laundry	4	fan	4	mayonnaise	6
zipper	4	pancake	4	grandpa	4	mermaid	4	sunflower	4	panties	4	strong	2
this	1	princess	3	talk	1	zero	4	railroad	4	wild	3	rest	2
lake	2	hamster	4	ghost	4	please	1	lap	4	globe	4	lazy	2
shower	4	dragonfly	4	fight	1	tear	2	trick	2	moose	4	pasta	12
jeep	4	away	1	build	1	trombone	4	calendar	4	beak	4	elf	4
tape	2	surfboard	4	why	1	vacuum	8	submarine	4	eyelash	4	bandage	4
nap	3	buffalo	3	smell	2	seaweed	4	paintbrush	4	pirate	4	bang	2
cowboy	2	pebble	4	bathe	4	underpants	4	cent	3	happen	2	cliff	3
band	3	find	1	classroom	4	parachute	4	tickle	4	escalator	4	whisker	3
guess	1	marble	4	favorite	3	plum	4	somersault	4	skateboard	4	frown	3
funny	4	muffin	4	spinach	4	snowball	4	carrier	6	photograph	4	avocado	8
wave	2	angry	4	sweatshirt	4	dozen	4	footprint	4	ham	4	tag	3
stripe	4	cough	4	magnet	4	microphone	4	lick	2	porch	4	dairy	4
mosquito	4	bead	3	crab	4	now	1	bush	3	microscope	4	chew	3
soft	2	onion	4	queen	3	get	1	porcupine	3	dive	3	ugly	2
not	1	acorn	4	clay	4	walrus	6	dresser	6	too	1	drip	4
ostrich	4	potato	3	syrup	6	stomach	3	forehead	4	let	1	stroller	4
dream	2	bathrobe	4	scorpion	8	dolphin	10	woodpecker	4	part	2		
pilot	3	pool	2	grocery	3	cub	4	volleyball	4	toad	4		
gas	3	nickel	3	party	1	neat	3	trailer	4	sandal	4		
merry-go-round	4	fin	4	dancer	4	flour	3	hose	3	raisin	4		