Computational Linguistics and Chinese Language Processing
Vol. 14, No. 1, March 2009, pp. 45-84

Assessing Text Readability Using Hierarchical Lexical Relations Retrieved from WordNet

Shu-Yen Lin*, Cheng-Chao Su*, Yu-Da Lai*, Li-Chin Yang* and Shu-Kai Hsieh*

Abstract

Although some traditional readability formulas have shown high predictive validity in the *r* = 0.8 range and above (Chall & Dale, 1995), they are generally not based on genuine linguistic processing factors, but on statistical correlations (Crossley *et al.,* 2008). Improvement of readability assessment should focus on finding variables that truly represent the comprehensibility of text as well as the indices that accurately measure the correlations. In this study, we explore the hierarchical relations between lexical items based on the conceptual categories advanced from Prototype Theory (Rosch *et al.*, 1976). According to this theory and its development, basic level words like *guitar* represent the objects humans interact with most readily. They are acquired by children earlier than their superordinate words like *stringed instrument* and their subordinate words like *acoustic guitar*. Accordingly, the readability of a text is presumably associated with the ratio of basic level words it contains. WordNet (Fellbaum, 1998), a network of meaningfully related words, provides the best online open source database for studying such lexical relations. Our study shows that a basic level noun can be identified by its ratio of forming compounds (*e.g.* chair → armchair) and the length difference in relation to its hyponyms. We compared graded readings for American children and high school English readings for Taiwanese students by several readability formulas and in terms of basic level noun ratios (*i.e.* the number of basic level noun types divided by the number of noun types in a text). It is suggested that basic level noun ratios provide a robust and meaningful index of lexical complexity, which is directly associated with text readability.

* Department of English, National Taiwan Normal University
E-mail: {yenyenet, vennysu, yudalai, lchyang1112, shukai}@gmail.com

[Received January 1, 2009; Revised April 11, 2009; Accepted April 15, 2009]

Keywords: Readability, Prototype Theory, WordNet, Basic Level Words, Compounds.

1. Introduction

Traditional methods of measuring text readability typically rely on surface-level linguistic information such as the counting of sentences, words, syllables, or letters. Caution has long been taken in correlating these formulas with the reading process (Davison & Kantor, 1982; Rubin, 1985). In light of the many psycholinguistic findings on the reading process (Just & Carpenter, 1987; Perfetti, 1985; Rayner & Pollatsek, 1994), we start our research by assuming, in line with Rosch *et al.*'s Prototype Theory (Rosch & Mervis, 1975, Rosch *et al.*, 1976) and its later development (Rosch, 1977, 1978; Coleman & Kay, 1981; Lakoff, 1986; Tversky, 1990; Ungerer & Schmid, 1996), that words form conceptual hierarchies (*e.g.* furniture → chair → armchair) with lexical items at different levels posing varied processing difficulties. Putting the logic into templates, the measurement of the lexical difficulty of a text may be done by calculating the hierarchical levels at which its words fall. The best tool for our study is WordNet, a large, open source electronic lexical database of English, in which the different senses of words are interlinked in hierarchical structures by means of conceptual-semantic relations.

Our research was comprised of two stages. In the preliminary experiments, we utilized WordNet to identify the characteristics of basic level nouns. It was found that a basic level noun can be identified by its ratio of forming compounds (*e.g.* chair → armchair) and the length difference in relation to its full hyponyms. In the subsequent experiment, we compared selected readings in terms of their basic level noun ratios and their values calculated by several readability formulas. It is shown that basic level noun ratios are highly correlated with the text levels. Our study also indicates that there is a basic level in a lexical hierarchy which is easier to comprehend than its upper or lower levels. This finding challenges the intuitive idea underlying McNamara *et al*. (2002) that a word having more hypernym levels is more concrete, thus, easier to comprehend, and fewer hypernym levels indicate more abstract language that is harder to understand.

The remainder of this paper is organized as follows: Section 2 reviews the common indices that form the base of many traditional readability formulas and the criticism they have received. In Section 3, we review Prototype Theory and discuss how it can aid us in finding the lexical difficulty of a text. Section 4 is about methodology – how to identify basic level words and how to assess the validity of our method against other readability formulas. Section 5 reports the results of the assessment and discusses the strength and weaknesses of our approach. In this section, we also suggest what can be done in subsequent research.

2. Literature Review

In this section we first summarize the indices of traditional readability formulas, give an account of the criticism these formulas meet, and introduce the purpose of our study. Among the multitude of factors underlying the reading process, we will focus on the lexical index.

2.1 Indices of Readability

2.1.1 Vocabulary Difficulty

The earliest work on readability measurement goes back to Thorndike (1921) where word frequency in a corpus is considered an important index in computing vocabulary complexity. This is based on the logic that the higher the frequency of a word, the more common and easier it is. Followers of this logic compiled word lists that include often-used and seldom-used words where the presence or absence of particular words on the lists assesses vocabulary difficulty, thus text difficulty.

Vocabulary difficulty is also measured by word length in many formulas, *e.g.*, the Flesch formula (Flesch, 1943, 1948, 1950) and FOG formula (McCallum & Peterson, 1982), or in terms of number of syllables (Fry, 1968). This is based on another intuitive assumption that the longer a word is, the more difficult it is to comprehend (Bailin & Grafstein, 2001).

2.1.2 Syntactic Difficulty

Syntactic complexity is another index in many readability formulas (Chall & Dale, 1995). For Dale & Chall (1948), Flesch (1948), and McCallum & Peterson (1982), syntactic complexity boils down to the average length of sentences in a text, although they vary in how they determine and utilize sentence length. The formula designed by Heilman, Collins-Thompson, Callan, & Eskenazi (2007) is a more recent example of this type. They propose that grammar-based predictions can be combined with vocabulary-based predictions to produce more accurate predictions of readability for both first and second language texts. They also suggest that language technologies must account for morphological features in languages which have a rich morphology, an issue relevant to grammatical features.

Also taking account of syntactic complexity, Miltsakaki & Troutt (2007) bases their algorithm on three readability formulas: Lix, Rix, and Coleman-Liau. The number of sentences, words, long words (seven or more characters), and letters in the text are taken into account. Another example is Das & Roychoudhury's work (2006), which built a readability index for Bangla using average sentence length (total words/ total sentences) and number of syllables per 100 words (total syllables/ total words*100).

2.1.3 Semantic Difficulty

Semantic factors such as counting abstract words (Flesch, 1943; Cohen, 1975) and propositional density and inferences (Kintsch, 1974) have also been put into regression analyses of readability assessment. In addition to these projects, Wiener *et al*. (1990) proposes a scale based on ten categories of semantic relations, *e.g*., temporal ordering and causality, for assessing the utterance complexity. The reliability of the semantic scale was confirmed when it was applied to compare the utterances of fourth-, sixth-, and eighth-grade children, where significant differences in semantic density were found on their scale.

Since 1920, more than fifty readability formulas have been proposed in the hopes of providing tools to measure readability more accurately and efficaciously (Crossley *et al*., 2007). Nonetheless, it is not surprising to see criticism over these formulas, given that reading is an extremely complex process.

2.2 Criticism of the Traditional Readability Formulas

Although classic readability formulas provide a quick and easy method of predicting readability, they are often criticized for being superficial, unstable, or unable to offer information about deeper levels of text processing (McNamara *et al*., 1996).

2.2.1 Criticism of Lexical Difficulty Measurement

Bailin & Grafstein (2001) question the validity of measuring vocabulary difficulty by the number of syllables per word or by the presence of words in a word list. They question the legitimacy of assessing vocabulary difficulty in terms of word length by showing that many mono- or bi-syllabic words are actually more esoteric, *i.e*. more unfamiliar, than longer polysyllabic terms. They also argue that the proposed link between readability and a vocabulary list of word frequency is narrowly based on the prerequisite that words in a language remain stable. The prerequisite, however, seems implausible as different socio-cultural groups have different core vocabularies and rapid cultural change makes many words out of fashion.

2.2.2 Criticism of Syntactic Difficulty Measurement

Bailin & Grafstein (2001) also point out the flaw of a simple equation between syntactic complexity and sentence length by giving the sample sentences as follows:

(1) I couldn’t answer your e-mail. There was a power outage.

(2) I couldn’t answer your e-mail because there was a power outage.

In terms of both absolute length and number of words, (2) is longer than (1), thus computed as more difficult by traditional readability formulas. Nevertheless, the subordinator "because" in (2), which explicitly links the author's inability to e-mail to the power outage, actually aids comprehension. As such, the authors suggest that language-oriented criteria be proposed, including deviations from prescriptive grammar, style (relative clauses, garden-path phrases, left-branching structures, *etc.*), and required background knowledge.

2.2.3 Criticism of Statistical Legitimacy

The correlation between the indices and the measured variables was also challenged from the viewpoint of statistical legitimacy. Hua & Wang (2007) point out the methodological issue in the creation of the traditional readability formulas. The typical initial step is to select, as the criterion passages, standard graded texts whose readability has already been agreed upon. The next step is to sort out the factors that may affect the readability of the text. The factors that are highly correlated with the text difficulty are chosen as independent variables in regression analysis for forming a readability formula. The researchers, however, did not ascertain whether the factors incorporated into their regression model actually have a cause-effect relationship with the dependent variable, *i.e.*, readability. Word length, used to equate semantic complexity, and sentence length, used for syntactic complexity, are intuitively correlated with readability, but non-scientifically correlated. Therefore, the authors suggest that researchers first analyze the independent variables qualitatively to confirm their cause-effect relationship with readability.

Challenge also goes to the selection of criterion passages. Schriver (2000) suggests that readability formulas are inherently unreliable because they depend on criterion passages too short to reflect cohesiveness, too varied to support between-formula comparisons, and too text-oriented to account for the effects of lists, enumerated sequences, and tables on text comprehension.

2.3 Purpose of Research

The criticisms of the traditional readability formulas by the various authors have a lot in common. They all urge adoption of language-oriented criteria based on independent evidence and a closer re-examination of the genuine relationship between the variables and the texts. It is our belief that this can only be done if we take account of the deeper levels of text processing. Reading is a multidimensional process; our pilot study aims to examine how a reader interacts with a text at the **lexical** level. We propose that the hierarchical status of a lexical item in our mental lexicon is a possible factor that affects lexical comprehensibility. We further suggest that there is a basic level in the lexical hierarchy which is the easiest to comprehend and serves as a meaningful indicator of text readability. To that end, we resort to Prototype Theory, which was proposed and developed by Rosch *et al*. (1976), among others.

3. Prototype Theory and Lexical Difficulties

3.1 Prototype Theory

Prototype Theory was brought to cognitive linguistics by Rosch *et al.* (1976). The notion of *prototype* can be understood in two ways. First, prototype is used either to refer to object members that first come to one's mind in an association experiment, or to those that can be recognized faster than other category members in a verification task. For example, when asked to give an example of "bird", "robin" is more frequently cited than "ostrich". Various researchers (Rosch, 1978; Lakoff, 1986; Brown, 1990; Tversky, 1990) use different names to label the prototypical member – "best example of a category", "salient examples", "clearest cases of category membership", or "central and typical members." The other way to define prototype is from a genuinely cognitive viewpoint. Prototype can be viewed as a mental representation, specifically as some kind of cognitive reference point (Rosch & Mervis, 1975; Coleman & Kay, 1981; Lakoff, 1986). Taking the two viewpoints together, we can view prototype as the central member or the cognitive reference point which other members of the category are anchored to. Through the anchoring process, cognitive categories are formed.

The members within a particular cognitive category are anchored to the prototype with different parameters – whether the members are perceived as *gestalt*, how many category-wide attributes are shared by the members, and how homogeneous or heterogeneous the members are. The representation of a bird, for instance, does not consist of a set of features that all birds have. A robin or a penguin as a category member of the bird is anchored to the most typical or ideal category member of the bird (which may not exist in real life). Since a robin shares more of the features characteristic of a prototypical bird than a penguin shares, it is usually viewed by subjects as a better example of a bird.

The same mental anchoring process can be applied to broader human categorization of those readily identifiable organisms and objects that surround us (Ungerer & Schmid, 1996: 60). As a result, entities within the cognitive category of DOG can be categorized as a "dog", a "terrier", a "Scottish terrier", a "mammal", or an "animal". These cognitive categories are connected with each other in a hierarchical pattern. In this example, if we look at their relationship from the "bottom" of the hierarchy, Scottish terriers are subordinate to terriers, and terriers are subordinate to dogs. If we look at them from the "top" of the hierarchy, animals are viewed as superordinate to mammals, and mammals as superordinate to dogs.

Turning to early interpretations of the *basic level* from the psychological viewpoints by Brown (1958) and Kay (1971), the basic level is where human beings perceive the most obvious difference between the organisms and objects of the world. Imagine an everyday conversation where a person says "Who moved that piano?" The naming of an object with "piano" will not strike us as noteworthy until the alternative "Who moved that stringed instrument?" is brought to our attention. Both terms are truth-conditionally adequate, but only the former is commonly used. The superordinate word "stringed instrument" is not used because its meaning encompasses many basic level words, *i.e*. many kinds of musical instruments. In our example, using the word "stringed instrument" is too vague to represent the object: "Stringed instrument" does not, as "piano" does, denote the most obvious difference of the object from the other objects in the world. Likewise, using a subordinate level word, *e.g*. a "grand piano", on a similar occasion is unusual except when the differentiation between different types of pianos is required.

In ranking typicality of objects, the basic level is where the largest bundles of naturally correlated attributes are available for categorization (Rosch *et al.*, 1976; Ungerer & Schmid, 1996: 67). In addition, the basic level is where gestalt perception occurs to the greatest extent, and this is particularly easy for prototypical examples. An "apple" has reddish or greenish skin, white pulp, and a round shape, while it is hard to pinpoint the features of "fruit". For a layman, hardly any significant features can be added to "crab apple".

The underlying cognitive anchoring process of the psychological reality of the basic level and the prototype are very similar. In the same way as other peripheral members of a category are anchored to the prototypical member, other non-basic levels, namely *superordinate* and *subordinate levels*, are anchored to the basic level. Ungerer and Schmid (1996: 72) point out the two are actually a kind of symbiosis underpinned by two interdependent principles: First, prototype categories are most fully developed on the basic level. Second, basic levels only function as they do because they are structured as prototypical categories.

The first principle can be explained by our earlier discussion on the basic level. Recall that this level offers the largest amount of correlated attributes, and the attributes are accumulated in their most completed form in the prototype and expressed by the category name (*e.g*., "Robin", as the typical example of the category BIRD, accumulates most correlated attributes of medium size, feathers, flying and singing ability, *etc*. in a complete form.)

As for the second principle, maximization of the efficiency of basic level categories by prototypes can be used to explain it (Rosch, 1977, 1978). That is, prototypes maximize the discontinuities or the distinctiveness of the basic level categories as they induce not only the greatest number of attributes shared inside the category, but the greatest number of attributes not shared by members of other categories. A typical example of a bird like "robin" can, while a non-typical example of the bird like "penguin" cannot, be easily distinguished from the category of fish because the latter shares more attributes with fish.

Developmentally, basic level categories are acquired earlier by children than their superordinate and subordinate words. Conceptually, the basic level category represents the concepts humans interact with most readily. Applying the hierarchical structure of conceptual categorization to lexical comprehensibility, we suggest that a concept at the basic level, hence, the word that denotes the concept, which we call a basic level word, is easier for the reader than its superordinate and subordinate words. If this is correct, then one text should be easier than another if it contains more basic level words. As the three-leveled hierarchy refers specifically to nouns in Prototype Theory, we confine our current study to the nominal category only. The best tool to study the relevant hierarchical relations in a broad framework with computational techniques is WordNet.

3.2 WordNet – A Hierarchically-Structured Lexical Database of English

WordNet is a large online electronic lexical database of English. The words are interlinked by means of conceptual-semantic and lexical relations. Its underlying design principle has much in common with the hierarchical structure proposed in Prototype Theory. In the vertical dimension, the hypernym/hyponym relations among the nouns can be interpreted as hierarchical relations between conceptual categories. For instance, the direct hypernym of "apple" in WordNet is "edible fruit". One of the direct hyponyms of "apple" is "crab apple". Note, however, that hypernyms and hyponyms are relativized notions in WordNet. Theoretically speaking, any word may have hypernyms and hyponyms. "Crab apple," being a hyponym of "apple," is also a hypernym in relation to "Siberian crab apple". An ontological tree may well exceed three levels. There are no labels or ready-to-be-used statistical information in WordNet that tell us which nouns fall into the basic level category. In the following sections we try to retrieve the basic level nouns as defined in Prototype Theory and apply the results in assessing text readability.

4. Methodology

Three experiments were conducted. In the first experiment, we utilized the nouns used in Rosch *et al*'s experiments in order to discover their quantitative properties. The second experiment followed up the first one and tried to pinpoint the criteria of determining the quality of the nouns being basic. In the third experiment, we computed and compared the basic level noun ratios and readability scores of graded readings. Our results indicate that basic level noun ratios provide a robust and meaningful index of text readability.

4.1 Experiment 1

4.1.1 Design of Experiment 1

In our initial experiment, we examined the eighteen basic level words identified by Rosch *et al*. (1976: 388), checking their word length, lexical complexity, and their direct hypernyms as well as direct hyponyms in WordNet. We speculate that a basic level word has these features: (1) It is relatively short (containing fewer letters than its hypernyms/hyponyms on average); (2) It is morphologically simple[1]; (3) It has more direct hyponym synsets than direct hypernym synsets[2]. Notice that some entries in WordNet are made up of more than one word. We assume that an item composed of two or more words is NOT a basic level word. A lexical entry composed of two or more words is a compound. The first word of a compound noun may or may not be a noun, and there may or may not be spaces or hyphens between the component words of a compound. For words having more than one sense, we focused only on the sense occurring in Rosch *et al*'s experiment. As an example, the noun "table" has six senses (*i.e.* synsets) in WordNet, but only the information in the sense of "a piece of furniture" is computed. Table 1 summarizes the results of Experiment 1.

Table 1. Eighteen basic level words in comparison with their direct hypernyms and direct hyponyms on word length, number of synsets, and morphological complexity*

Target word	Index of the inquired synset	Basic level		Direct hypernym			Direct hyponym		
		Word length	Morph. Complexity	Average word length	Number of synsets	Morph. Complexity	Average word length	Number of synsets	Morph. Complexity
guitar	0	6	A	18	1	B	8.8	6	A, B
piano	0	5	A	19	3	B	9.6	3	A, B
drum	0	4	A	20	1	B	7.4	8	A, B
apple	0	5	A	8.3	2	A, B	10.6	3	A, B
peach	2	5	A	8.7	2	B	N/A	N/A	N/A
grape	0	5	A	11	1	B	11.8	3	A, B
hammer	1	6	A	8	1	B	9.7	7	A, B
saw	1	2	A	8	1	B	8.7	7	A, B

[1] It has been noticed by Ungerer & Schmid (1996: 98) that basic level words tend to be short and mono-morphemic words, and that the superordinate and subordinate words tend to be longer and morphologically complex words. Nevertheless, to our knowledge, no systematic studies have been done prior to ours on these quantitative features.

[2] Both hyponyms and hypernyms are grouped into synsets in WordNet. A synset is a set of synonyms. The direct hyponyms of "guitar", for instance, are grouped into six synsets: (1) acoustic guitar, (2) bass guitar, (3) cittern, cithern, cither, citole, gittern, (4) electric guitar, (5) Hawaiian guitar, steel guitar, and (6) uke, ukulele. In contrast, "guitar" has only one direct hypernym synset, *i.e.* "stringed instrument". Accordingly, the number of direct hypernym synsets of "guitar" is 1, and the number of its direct hyponym synsets is 6, as shown in Table 1.

screwdriver	0	11	B	8	1	B	19.8	3	B
pant	1	4	A	7	1	A	8.9	18	A, B
sock	0	4	A	5.5	1	A	7.8	5	A, B
shirt	0	5	A	7	1	A	8.1	9	A, B
table	1	5	A	14.3	1	A	10.4	26	A, B
lamp	1	4	A	14.3	1	B	9.7	3	A, B
chair	0	5	A	4	1	A	10.7	15	A, B
car	0	3	A	14.5	1	A, B	8.3	31	B
bus	0	3	A	15	1	A, B	10.8	3	B
truck	0	5	A	14.5	1	A, B	8.7	11	B

*A refers to “single word”. B refers to “compound”.

4.1.2 Results of Experiment 1

The results confirm our prediction. First, the average word length (number of letters) of both the hypernyms and the hyponyms is much longer than that of the basic level words. Although many researchers have pointed out that absolute word length is not a meaningful indicator of lexical complexity (Davidson & Kantor, 1982; Bailin & Grafstein, 2001; Hua & Wang, 2007), to our knowledge no researchers so far have been able to propose a better algorithm to account for word length or be able to refute the intuition that word length plays a certain role in reflecting lexical difficulty.

Our results indicate that word length should be viewed in a relative sense; namely, there is a level in the lexical hierarchy that has the shortest word length in comparison with its higher and lower levels on average. That absolute word length is not a good index is manifested by the fact that in Table 1 some of the direct hypernyms are shorter than six letters in their average length, while some basic level words have only six letters. We have, however, observed consistency in the word length difference between levels of words within a hierarchy: The basic level words always contain the fewest number of letters[3]. The tendency is particularly strong when we compare the length of basic level words with the average length of their direct hyponyms[4].

---

[3] The tendency has only one exception, *i.e.* the direct hypernym of “chair” is shorter than “chair”.

[4] The word “screwdriver” seems to be an exception to the pattern we describe, as the average length of its direct hypernyms is shorter than the target word itself (8 < 11). Nevertheless, since “screwdriver” is a compound, *i.e.* composed of “screw” and “driver”, it is actually excluded by our basic level word criteria.

Our second finding from Experiment 1 is that these basic level words have many more direct hyponym synsets than direct hypernym synsets. Finally, in contrast to the basic level words which are morphologically simple, the direct hypernyms and the direct hyponyms are more complex. Many of the hypernyms are compounds. The hyponyms are even more complex. Every basic level word (except for "peach") has compounded hyponyms.

4.2 Experiment 2

4.2.1 Design of Experiment 2

Our first findings brought our attention to the relative length difference of the words at different levels and the disparity of their morphological structure. In particular, we found that basic level words display sharper contrast with their hyponyms than with their hypernyms, both in terms of word length difference and morphological structure complexity.

In this experiment, we set out to compute the difference between the length of the basic level words in Experiment 1 and the average length of their full hyponyms. We also examined the distribution of the compounds formed by the three levels of words - basic level words, their hypernyms, as well as their hyponyms. Additionally, we randomly came up with seven more words that appear to fall into the basic level category defined by Rosch *et al*. (1976). The results of this experiment are shown in Table 2.

Table 2. The twenty-five basic level words – Word length differences, compound ratios, and distribution of compounds

Hypernym / Basic level word / Hyponym	Index of the in-quired synset	Average hypo-nym l. – target word l.	# of com-pounds / # of full hypo-nyms	Cpd ratio (%)	Number of compounds at hyponymous levels					
					1st level	***2nd level***	***3rd level***	***4th level***	***5th level***	***6th level***
stringed instrument	0	- 10.2	1 / 85	1	1	0	0	0		
guitar	0	**2.8**	5 / 12	**42**	5					
acoustic guitar	0	N/A	N/A	N/A						
keyboard instrument	0	- 9.0	0 / 35	0	0	0	0			
piano	0	**6.0**	8 / 16	**50**	4	4				
grand piano	0	1.9	3 / 8	38	3					
baby grand piano	0	N/A	N/A	N/A						

percussion instrument	0	- 12.5	0 / 68	0	0	0	0			
drum	0	**3.4**	5 / 14	**36**	5					
bass drum	0	N/A	N/A	N/A						
edible fruit	0	- 3.1	0 / 258	0	0	0	0	0		
apple	0	**5.5**	5 / 29	**17**	5	0	0			
crab apple	0	3.6	2 / 8	25	2					
Siberian crab	0	N/A	N/A	N/A						
edible fruit	0	- 3.1	0 / 258	0	0	0	0	0		
peach	2	**N/A**	N/A	**N/A**						
N/A	N/A	N/A	N/A	N/A						
edible fruit	0	- 3.1	0 / 258	0	0	0	0	0		
grape	0	**4.5**	6 / 17	**35**	3	2	1			
muscadine	0	N/A	N/A	N/A						
hand tool	0	1.0	0 / 217	0	0	0	0	0		
hammer	1	**3.9**	7 / 16	**44**	7	0				
ball-peen hammer	0	N/A	N/A	N/A						
hand tool	0	1.0	0 / 217	0	0	0	0	0	0	
saw	1	**5.7**	25 / 30	**83**	13	12	0			
bill	7	N/A	N/A	N/A						
hand tool	0	1.0	0 / 217	0	0	0	0	0	0	
screwdriver	0	**8.8**	4 / 4	**100**	4					
flat tip screwdriver	0	N/A	N/A	N/A						
garment	0	1.0	4 / 306	1	3	1	0	0	0	
pant	1	**4.9**	10 / 49	**20**	9	1				
bellbottom trousers	0	N/A	N/A	N/A						
hosiery	0	1.7	0 / 28	0	0	0				
sock	0	**3.8**	5 / 13	**38**	5					
anklet	0	N/A	N/A	N/A						
garment	0	1.0	4 / 306	1	3	1	0	0	0	
shirt	0	**3.2**	8 / 17	**47**	8	0				

camise	0	N/A	N/A	N/A						
furniture	0	0.0	4 / 244	2	4	0	0	0	0	
table	1	**5.1**	36 / 77	**47**	29	7	0	0		
altar	0	N/A	N/A	N/A						
source of illumination	0	- 11.9	0 / 107	0	0	0	0	0	0	
lamp	1	**6.0**	4 / 4	**100**	3	1				
Aladdin"s lamp	0	N/A	N/A	N/A						
seat	2	5.0	7 / 101	6.9	3	3	1	0		
chair	0	**5.6**	31 / 48	**65**	17	14	0			
armchair	0	2.7	0 / 10	0	0	0				
captain's chair	0	N/A	N/A	N/A						
motor vehicle	0	- 4.3	0 / 151	0	0	0	0	0		
car	0	**5.3**	21 / 75	**28**	19	2				
amphibian	0	1.0	0 / 2	0	0					
public transport	0	- 6.7	0 / 38	0	0	0	0			
bus	0	**7.8**	3 / 5	**60**	3					
minibus	0	N/A	N/A	N/A						
motor vehicle	0	- 4.3	0 / 151	0	0	0	0	0		
truck	0	**4.5**	15 / 48	**31**	10	5	0			
dump truck	0	N/A	N/A	N/A						
canine	1	4.3	0 / 287	0	0	0	0	0	0	0
dog	0	**8.0**	50 / 235	**21**	11	21	16	2	0	
puppy	0	N/A	N/A	N/A						
feline	0	3.1	0 / 123	0	0	0	0			
cat	0	**6.1**	35 / 87	**40**	4	30	1			
domestic cat	0	- 4.0	0 / 32	0	0					
kitty	2	N/A	N/A	N/A						
publication	0	- 1.3	2 / 192	1	0	1	0	0	0	
book	0	**6.5**	38 / 139	**27**	17	11	7	3	0	
authority	6	- 1.0	0 / 1	0	0					
last word	0	N/A	N/A	N/A						

language unit	0	- 3.0	0 / 290	0	0	0	0	0	0	0
word	0	**6.5**	35 / 185	**19**	28	7	0	0	0	
anagram	6	1.0	0 / 1	0	0					
antigram	0	N/A	N/A	N/A						
material	0	1.1	17 / 591	2.9	15	2	0	0		
paper	0	**4.1**	59 / 173	**34**	40	18	1			
card	0	3.1	14 / 57	25	6	8				
playing card	0	- 5.4	0 / 49	0						
movable barrier	0	- 6.2	0 / 44	0	0	0	0			
door	0	**5.8**	18 / 23	**78**	13	5				
car door	0	3.0	0 / 2	0	0					
hatchback	0	N/A	N/A	N/A						
leaf	1	4.7	2 / 21	10	2	0	0			
page	0	**5.0**	5 / 18	**28**	5	0				
full page	0	N/A	N/A	N/A						

In the first column, the basic level words (*e.g*. "guitar") are boldfaced, with the (or one of the) direct hypernym(s) (*e.g*. "stringed instrument") given above and its first-occurring direct hyponym (*e.g*. "acoustic guitar") placed under it. When the basic level word has more than one level of hyponym, the first word occurring at the second hyponymous level was also examined such that the word "hatchback" was under "car door".

As in Experiment 1, with respect to words having more than one sense, we focused only on the sense defined in Rosch *et al*. (1976). As an example, the noun "table" has six senses (*i.e*. synsets) in WordNet, but only the information in the sense of "a piece of furniture" is computed. Which synset conforms to the sense in Rosch *et al.* (1976) was decided manually. In WordNet, each sense of a word is indexed numerically. We put the index of the inquired synset in the second column. Notice that the first sense of a word has the number 0, and the second sense, the number 1, and so on and so forth.

All other information was retrieved by a program we wrote based on NLTK-0.9.5, which was downloaded at http://www.nltk.org/. (NLTK had been updated to version 0.9.9 by the time of revision of this paper.) Our own program can be downloaded at http://lope.eng.ntnu.edu.tw/lopedia/index.php/Image:Compound_ratios_and_word_length_difference_in_WordNet.doc#filelinks). We set the hyponym depth in our program at 100 levels.

The third column "Full hyponym length minus target word length" computes the difference between the length of the target word and the average length of its full hyponyms.

The word "stringed instrument" has, for example, 85 hyponyms in total with their average length being 8.79 letters. The length difference is 8.79 minus 19, which equals -10.2. A negative value in this column thus means that the target word is longer than the average length of its full hyponyms. On the other hand, a positive value conveys longer average hyponym length than the target word.

The fourth column computes the ratios of compounds composed of the target word in the full hyponyms. Our program searches the full hyponyms for compounds that are formed by the target words[5]. Such a compound may end with the target word, which constitutes the major compounding pattern we have observed, *e.g*. "school bus" is a compound hyponym of "bus". The other way to form a compound hyponym is to start with the target word. The only examples we know, however, are "icefall," "ice pack," and "ice shelf" in the second synset of "ice". In light of the existence of compounds like those, we also include this compounding template in our program. The program virtually searches for compounds that contain the target word, assuming that the target word may occur in the front, middle, or end position of the compound.

As an example of the compounding behavior and the computation of compound ratio, among the twelve (full) hyponyms of "guitar," five are compounds formed by "guitar". They are "acoustic guitar," "bass guitar," "electric guitar," "Hawaiian guitar," and "steel guitar". The compound ratio of "guitar" is accordingly 5/12 = 42%. By contrast, only one hyponym of the full eighty-five hyponyms of "stringed instrument" is a compound containing "stringed instrument" (*i.e*. "bowed stringed instrument"), and its compound ratio is 1/85 = 1%. As for "acoustic guitar," it has no hyponyms. These compound ratios are given in the fifth column.

We also keep record of the levels where compounds occur, which we display in the rightmost columns.

4.2.2 Results of Experiment 2

Several regular patterns can be observed in Table 2. In terms of word length difference, the basic level words show the greatest positive values across the board. Each basic level word enjoys greater length difference than its direct hypernym and its first-occurring hyponym do. On the other end of the length difference spectrum are the direct hypernyms. Of the twenty-five direct hypernyms, ten have negative values. Only four have positive length differences greater than 3. This is likely to result from the fact that the direct hypernyms in Table 2 are mostly compound words, thus, tend to be long. If we go one level higher, the length difference may decrease. For example, the direct hypernym of "edible fruit" is "fruit",

[5] When a word is compounded with another word to form a compound, this compounded word becomes a hyponym of the target word and is unlikely to become a hypernym of the target word.

whose length is much shorter.

The most significant finding from Experiment 2 is that basic level words have the highest compound ratios. In comparison with their hypernyms and hyponyms, they are much more frequently used to form compounds. Although some hyponyms like “grand piano” and “crab apple” also have high compound ratios, they should not be taken as basic level items because these compounds often contain the basic level words themselves (*e.g.* “Southern crab apple” contains “apple”), indicating that the ability to form compounds is actually inherited from the basic level words.

Our data pose a challenge to Prototype Theory in that a subordinate word of a basic level word may act as a basic level word itself. The word “card,” a hyponym of “paper,” is of this type. With its high compound ratio of 25%, “card” may also be deemed to be a basic level word. This fact raises another question as to whether a superordinate word may act as a basic level word as well. One might, however, object that it is doubtful whether “card” is really subordinate to “paper” in the framework of Prototype Theory. That is to say, it takes independent evidence to prove that the hyponyms of these twenty-five basic level words in WordNet correspond to the subordinate words defined by Prototype Theory and that the hypernyms correspond to the superordinate words. We leave this issue aside for reasons that will be clear when we describe the design of the next experiment. At this moment, suffice it to say that the way we identify basic level words in WordNet is not based on how many levels of hyponyms or hypernyms a word has or on which specific level in the hierarchy a word falls.

Many of the basic level words in Table 2 have three or more levels of hyponyms. This indicates that what is cognitively basic may not be low in the ontological tree. A closer look at the distribution of the compounds across the hyponymous levels reveals another interesting pattern. Basic level words have the ability to permeate two to three levels of hyponyms in forming compounds. In contrast, words at their hypernymous levels do not have such ability, and their compounds mostly occur at their direct hyponymous levels only. Words at their hyponymous levels rarely, if ever, form compounds.

4.3 Experiment 3

4.3.1 Design of Experiment 3

The goal of this experiment is to show that whether a word belongs to the basic level affects its comprehensibility, which in turn affects the readability of a text. If this is correct, when all other factors are equal, an easy text should contain more basic level words than a difficult text. Put in fractional terms, we attempt to show that the proportion of basic level words in a text is correlated with the readability of the text.

To achieve this goal, we need independent readability samples to compare with our prediction. Nevertheless, as readability is a subjective judgment that may vary from one person to another, such independent samples are extremely difficult, if possible, to obtain. In this study, we resorted to a pragmatic practice by selecting the online graded readings for American children and texts in English textbooks for senior high school students in Taiwan. Five open source readings ranging from grade one to twelve from edHelper.com (http://www.edhelper.com/ReadingComprehension.htm) were randomly selected for Experiment 3. Three textbooks from Sanmin Publishing Co., each used in the first semester of a different school year, were also used for this experiment. We tried to choose the same type of text, so that text type would not act as noise. All three high school English texts are informational. Due to the great divergence between the levels of children's readings, however, it was not easy to be strict with text types. Furthermore, since we do not have the facilities to run large-scale experiments yet, we limited the scope to around four-hundred-word texts at the Taiwanese high school level, and approximately two-hundred-and-fifty word texts at the American children's level. All selected readings are appended in Appendix A and B. Using the same program as in Experiment 2, we searched WordNet 3.0 for all the nouns occurring in these texts, except for proper names and pronouns. We referred only to the words in the particular sense occurring in the selected readings. We know that this practice, if used in a large-scale study, is applicable only if sense tagging is available.

Based on the results of the two preliminary experiments, we argue that the basic level noun index includes at least the following two quantitative features: (1) A basic level noun has great ability to form compounded hyponyms; (2) The length of a basic level noun is shorter than the average word length of its full hyponyms. These characteristics can be further simplified as the **Filter Condition** to pick out basic level nouns:

(1) Compound ratio (*i.e.* the number of the hyponyms which contain the target word divided by the number of the target word's full hyponyms) ≧ 20%;

(2) Length difference (*i.e.* the average length of the target word's full hyponyms minus the length of the target word) ≧ 2.

We set the compound ratio threshold at twenty percent for the following reasons. On the one hand, the compound ratios of all the basic level words in Experiment 2, except for "apple," "pant," and "word," are higher than twenty-five percent. On the other hand, these basic level words are derived from the psycholinguistic experiments by Rosch *et al.* (1976) which were designed to markedly manifest the human conceptual structure of categorization. Due to the special purpose of their experiments, these words are supposed to be the most typical basic level words in the English vocabulary. The quantitative data obtained using these words should be fine-tuned to a lower level to capture the representativeness of the other not-so-typical basic level words. As the compound ratios of "apple," "pant," and "word" are all near twenty percent, we approximate the threshold of the compound ratio to be twenty percent. In our future research, with more training data, the threshold will be further weighted.

The reason for setting up the condition of word length difference at two letters is the same as for the setting of the compound ratio: The basic level words in Experiment 2 are the most typical, therefore, on the upper end of the spectrum. Even though nineteen of the twenty-five words are shorter than their average hyponym by at least four letters, we set the word length difference condition at two letters. This threshold should also be further weighted in future research.

Note in passing that the second criterion differs fundamentally from the commonly used criterion of word length. Ours compares the target word with the average length of its full hyponyms. In our study word length is measured in relative terms: The word length difference is an index, not the word length itself.

Based on the two criteria of our filter condition, the information for each noun we need include the length of the target word, the average word length of its full hyponyms, the number of its full hyponyms, and the number of compounds of the target word, *i.e.* how many hyponyms of the word are compounds formed by the word. All computed values for each noun in the selected readings can be found in Appendix C and D. Words that pass the filter condition are displayed in red color with their compound ratios and length differences being boldfaced.

The next step of Experiment 3 was to compute the basic level noun ratios of the selected readings. Basic level noun ratios were obtained by dividing the number of basic level noun types in a text by the number of all noun types of the text. For example, the easiest text in our randomly selected online readings for American children ("Wash Your Hands" in Appendix A) contains 21 noun types (excluding proper names and pronouns), and 12 of them (*i.e.* the red items in the first table of Appendix C) reach the basic level noun threshold. The basic level noun ratio of this reading is accordingly 12/21 = 57.1%. Using the online software *Readability Calculations* (http://www. micropowerandlight.com/rd.html), we also obtained the scores of these texts computed by several readability formulas. These scores were then compared with the basic level noun ratios. We report and discuss the results of Experiment 3 in the next section.

5. Results and Discussion

Table 3 shows the raw scores and z-scores of the American children's readings and the Taiwanese high school texts calculated in terms of basic level noun ratios and by several readability formulas. With respect to the children's texts, Level 1 is the easiest level and Level 12 is the hardest.

Table 3. Raw- and z-scores of children and high school English texts computed by basic level noun ratios and several readability formulas

Measurement	Score	**Level 1 ~ 2**	**Level 3 ~ 4**	**Level 4 ~ 6**	**Level 7 ~ 8**	**Level 9 ~ 12**	**Book 1 Lesson 2**	**Book 3 Lesson 1**	**Book 5 Lesson 1**
Basic Level Word Ratio (%)	Raw score	57.1	48.3	32.6	28.6	21.1	39.3	26.1	25.0
	Z-score	*1.78*	*1.08*	*-0.17*	*-0.49*	*-1.09*	*0.36*	*-0.69*	*-0.78*
Dale_Chall	Raw score	4.7	6.3	7.2	6.6	7.8	4.5	7.0	7.1
	Z-score	*-1.42*	*-0.08*	*0.67*	*0.17*	*1.17*	*-1.59*	*0.5*	*0.59*
Flesch Grade Level	Raw score	0.3	4	6	6.3	8.1	1.5	7.9	9.3
	Z-score	*-1.59*	*-0.44*	*0.18*	*0.27*	*0.83*	*-1.21*	*0.77*	*1.2*
FOG	Raw score	2.8	9.7	19.1	13.2	18	*-1.1*	*17*	*21.4*
	Z-score	*-1.59*	*-0.56*	*0.85*	*-0.03*	*0.69*	*-1.1*	*0.54*	*1.2*
Powers	Raw score	3.3	4.7	5.5	5.2	6	*3.8*	*6.0*	*6.3*
	Z-score	*-1.65*	*-0.37*	*0.37*	*0.09*	*0.83*	*-1.19*	*0.83*	*1.1*
SMOG	Raw score	3.9	7.5	9.8	9.1	10.1	*9.8*	*11.9*	*11.9*
	Z-score	*-2.06*	*-0.67*	*0.21*	*-0.06*	*0.33*	*0.21*	*1.02*	*1.02*
FORCAST	Raw score	6.4	9	9.4	9.1	11.2	*7.4*	*11.5*	*10.9*
	Z-score	*-1.63*	*-0.2*	*0.02*	*-0.14*	*1.01*	*-1.08*	*1.18*	*0.85*
Spache	Raw score	1.8	3.0	3.6	3.8	4.2	*2.2*	*3.6*	*4.5*
	Z-score	*-1.63*	*-0.36*	*0.28*	*0.49*	*0.92*	*-1.21*	*0.28*	*1.23*

Diagrammatically, it is clear in Figure 1 that the basic level noun ratios of the American children’s texts decrease inversely proportionally to the difficulty levels of the selected readings. Readability scores of the same texts calculated by the traditional formulas contain more ups and downs.

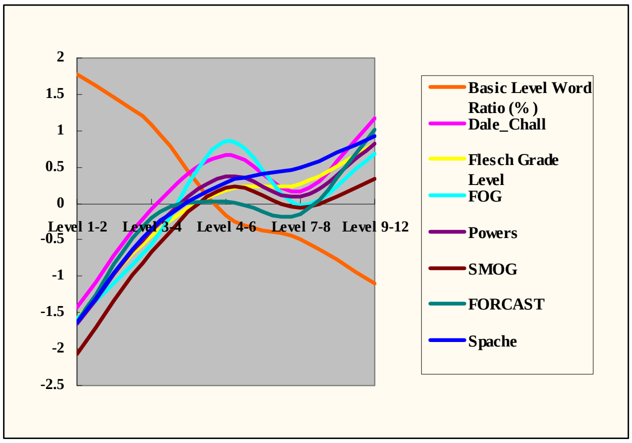

Figure 1. Readability of children's texts computed by basic level noun ratios and several readability formulas

As for the Taiwanese high school English texts, Figure 2 shows that Book 3, Lesson 1 and Book 5, Lesson 1 are rated similarly both in terms of basic level word ratios and by most of the readability formulas. We suspect that the textbooks are not well differentiated according to the levels of the students.

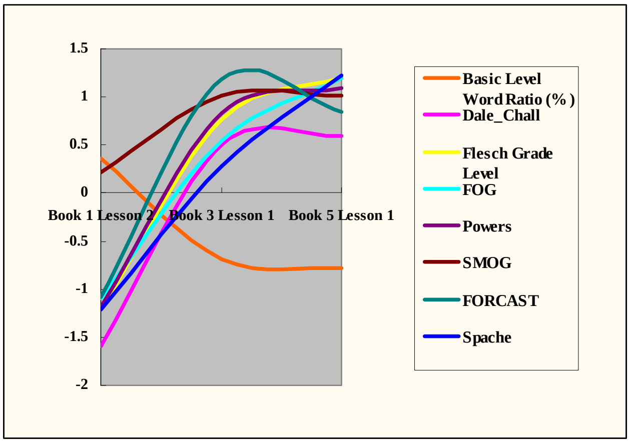

Figure 2. Readability of high school English texts computed by basic level noun ratios and several readability formulas

Overall, the basic level word ratios obtained in our study, both for high school students in Taiwan and for American children, conform to the levels of these texts, proving the usefulness of the basic level word concept in the assessment of readability.

The hierarchical relations in WordNet have also been utilized in Cohmetrix (McNamara *et al*. version 2.0), an online readability assessment software. It also uses the hierarchical relations as an index of conceptual difficulty:

> "A word having more hypernym levels is more concrete. A word with fewer hypernym levels is more abstract."

What the Cohmetrix calculates is the mean levels above the words (nouns, verbs, and adjectives). A word at the bottom of an ontological tree in WordNet is deemed in Cohmetrix to be the most concrete. Since what is more concrete is generally believed to be simpler, a word at the lowest level is viewed as the easiest conceptually. The higher the level reaches, the more abstract, hence, conceptually more challenging for the human processor it becomes. This may seem intuitively sound, but our study has clearly shown that the relations between lexical items in a hierarchy are not like a ladder, a metaphor that captures what the Cohmetrix calculation seems to imply of the relations between the lexical items in our mental lexicon. We used the online software of Cohmetrix and obtained the scores in Table 4.

Table 4. Raw- and z-scores of lexical conceptual difficulty computed by Cohmetrix

Measurement	Score	**Level 1 ~ 2**	**Level 3 ~ 4**	**Level 4 ~ 6**	**Level 7 ~ 8**	**Level 9 ~ 12**	**Book 1 Lesson 2**	**Book 3 Lesson 1**	**Book 5 Lesson 1**
Mean hypernym values of nouns	Raw score	5.721	5.637	4.831	5.55	4.623	4.935	4.65	4.823
	Z-score	*1.36*	*1.17*	*-0.58*	*0.99*	*-1.03*	*-0.35*	*-0.97*	*-0.59*

Table 4 shows that the mean hypernym values of the nouns in these texts are not correlated with the text levels. This is illustrated by the sharp up and down in Figure 3 and the big curve in Figure 4.

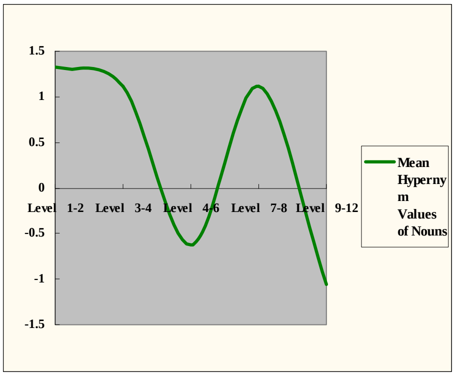

Figure 3. Readability of children's texts computed in terms of mean hypernym values of words

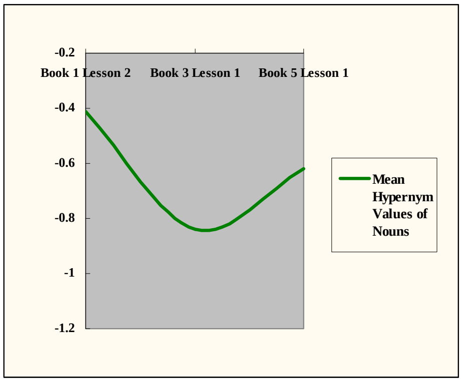

Figure 4. Readability of high school English texts computed in terms of mean hypernym values of words

These data suggest that a calculation that takes the lowest lexical hierarchical level as the most basic apparently misses the conceptual divide in our mental lexicon marked up by the basic level.

This paper is just the first step in assessing readability by lexical relations retrieved from WordNet based on conceptual categorization theories. Our study of readability assessment takes an approach that deviates remarkably from the traditional readability formulas. It looked for language-oriented variables that truly correlate with the reading process based on independent evidence. Prototype Theory, which is rooted in cognitive psychology and linguistics, gives us the idea that words form levels in our mental lexicon, with each level having its own characteristics and, accordingly, varied comprehensibility. It is our belief that these qualitative properties proposed in linguistic theories have corresponding quantitative features and retraceable distributions. The aim of our research is to find out these features by means of computational linguistic approaches and apply them in the assessment of text readability. The electronic lexical database of WordNet is an excellent tool to test our hypothesis. The results of our experiments are stimulating, but at the same time pose more challenges than achievements.

The filter condition of basic level nouns proposed in this study still leaves room to be fine-tuned and improved at least in two respects. First, the two criteria of compound ratios and word length difference have been used as sufficient conditions. More experiments will be designed for weighting these parameters in our future research. Specifically, we will study the distribution of compound words and the distribution of hyponyms over English words as one of our reviewers has pointed out that the distributions of these basic quantities affect the ratios. Furthermore, as another reviewer points out, the parameters must be able to be transformed into a scale in the future. Second, in addition to the lexical relations proposed in this study, there are presumably other relations between basic level words and their hypernyms/hyponyms that are retrievable via WordNet and other databases. These relations, if found, can further modify the basic level word criteria proposed in this study.

Doubts can be raised as to whether all basic level words are equally readable or easy. Can it be that some basic level words are in fact more difficult than others and some hypernyms/hyponyms of certain basic level words are actually easier than certain basic level words? Are basic level words frequent words in general? Can we substitute frequency for the quality of being basic if the two criteria have approximately the same indexing power? This question can be extended to whether the hierarchical relations between the lexical units in WordNet are correlated with word frequency, and if so, in what ways. We will try to answer these questions in a study of larger scale.

The examined words in this study are all nouns. Can we find relations between verbs, adjectives, and even adverbs like the hypernym/hyponym relations within the various levels of nouns? The tentative answer to this question is yes and no. Take the example of the verb "run". It has hypernyms in WordNet ("speed," "travel rapidly," *etc*.). It also has subordinate lexical relations called "troponym," which are similar to hyponyms of nouns. English verbs, admittedly, do not constitute compounds as often as English nouns, but other lexical relations may exist between the verbs, and the relations are likely to be retrievable.

As Bailin & Grafstein (2001) suggest, lexical difficulty assessment should take into account the socio-cultural groups whose core vocabulary and background knowledge differ considerably in specific fields. Our initial speculation in this respect is that every academic and professional discipline has its own set of basic level words. These words may be highly infrequent in everyday use of the language, but form the fundamental layer in the jargon in its own sphere. A truly useful readability measurement tool thus should correspond to the text category and meet the readers' personal needs.

Future readability assessment tools should also be able to report not only the difficulty levels of the texts according to the readers' background knowledge but also the difficulty itself. In the lexical dimension, the tool should highlight the high level vocabulary for the readers. An algorithm like our current application is working exactly in this direction.

Laying out the groundwork for further research, we aim to tackle the following issues as well. All traditional readability formulas implicitly suppose an isomorphic relation between form and meaning as if each word has the same meaning no matter where it occurs. We acknowledge that one of the biggest challenges of measuring readability is to disambiguate the various senses of a word in text as the same word may have highly divergent readability in different senses. Another tacit assumption made by the traditional readability formulas is that the units of all lexical items are single words. This assumption overlooks many compounds and fixed expressions, affecting the validity of these formulas. This raises the issue of segmentation. It is clear that the rating process applied in this study cannot be fully automated without successful segmentation of the text.

Although the small scale size of our experiments makes the validity of the results challengeable, its findings have provided the outlook of a large-scale project in the future. It has opened up a new approach to the assessment of text readability.

References

Bailin, A. & Grafstein, A. (2001). The linguistic assumptions underlying readability formulae: A critique. *Language and Communication*, 21(3), 285-301.

Brown, R. (1958). How shall a thing be called? *Psychological Review*, 65(1), 14-21.

Chall, J., & Dale, E. (1995). *Readability Revisited: The New Dale-Chall Readability Formula.* Cambridge, Massachusetts: Brookline Books.

Cohen, J. H. (1975). The effects of content are material on cloze test performance. *Journal of Reading*, 19(3), 247-250.

Coh-Metrix (Version 2.0) (Software). Memphis, TN: University of Memphis, Institute for Intelligent Systems. Available from http://cohmetrix.memphis.edu/cohmetrixpr/index.html

Coleman, L., & Kay, P. (1981). Prototype semantics: The English word "lie". *Language*, 57(1), 26-44.

Crossley, S. A., Dufty, D. F., McCarthy, P. M., & McNamara, D. S. (2007). Toward a new readability: A mixed model approach. In *Proceedings of the 29th Annual Conference of the Cognitive Science Society,* Austin, TX: Cognitive Science Society, 197-202.

Crossley, S. A., Greenfield, J., & McNamara, D. S. (2008). Assessing text readability using cognitively based indices. *Tesol Quarterly*, 42(3), 475-493.

Dale, E., & Chall, J. (1948). Formula for predicting readability. *Educational Research Bulletin* 27, 37-53.

Das, S., & Roychoudhury R. (2006). Readability modelling and comparison of one and two parametric fit: A case study in Bangla. *Journal of Quantitative Linguistics*, 13, 17-34.

Davison, A., & Kantor, R. (1982). On the failure of readability formulas to define readable texts: A case study from adaptations. *Reading Research Quarterly*, 17, 187-209.

edHelper. com, Reading Comprehensions, http://edhelper.com/ReadingComprehension.htm.

Fellbaum, C. (1998). *WordNet: An Electronic Lexical Database*. Cambridge, MA: The MIT press.

Flesch, R. (1943). *Marks of Readable Writing.* Ph.D. thesis.

—— (1948). A new readability yardstick. *Journal of Applied Psychology*, 32, 221-233.

—— (1950). Measuring the level of abstraction. *Journal of Applied Psychology*, 34, 384-390.

Fry, E. (1968). A readability formula that saves time. *Journal of Reading*, 11(7), 265-71.

Kay, P. (1971). Taxonomy and semantic contrast. *Language*, 47, 866-887.

Kintsch, W. (1974). *The Representation of Meaning in Memory*. Hillsdale, NJ: Erlbaum.

Heilman, M., Collins-Thompson, K., Callan, J., & Eskenazi, M. (2007). Combining lexical and grammatical features to improve readability measures for first and second language texts. In *Proceedings of the HLT/NAACL Annual Conference*, Rochester U.S.A., 460-467.

Hua, N. & Wang, G. (2007). Lun chuantong keduxing gongshi de bu-kexuexing. [On the non-scientific aspects of traditional readability formulae], *KaoShiZhouKan*, 18, 119-120.

Just, M. A., & Carpenter, P. A. (1987). A theory of reading: From eye fixations to comprehension. *Psychological Review*, 87, 329-354.

Lakoff, G. (1986). Classifiers as a reflection of mind. In C. G. Craig (Ed.), *Noun Classes and Categorization*, 13-52. Amsterdam: John Benjamins.

McCallum, D. R., & Peterson, J. L. (1982). Computer-based readability indices. In *Proceedings of the ACM '82 Conference*.

McNamara, D. S., Kintsch, E., Butler-Songer, N., & Kintsch, W. (1996). Are good texts always better? Interactions of text coherence, background knowledge, and levels of understanding in learning from text. *Cognition and Instruction*, 14, 1-43.

McNamara, D. S., Louwerse, M. M., & Graesser, A. C. (2002). Coh-Metrix: Automated cohesion and coherence scores to predict readability and facilitate comprehension. Unpublished technical report: University of Memphis.

Miltsakaki, E., & Troutt, A. (2007). Read-X: Automatic evaluation of reading difficulty of web text. In T. Bastiaens & S. Carliner (Eds.), *Proceedings of World Conference on E-Learning in Corporate, Government, Healthcare, and Higher Education 2007*, Chesapeake, VA: AACE, 7280-7286.

NLTK version 0.9.5 (Natural Language Toolkit) http://www.nltk.org/Home.

Perfetti, C. A. (1985). *Reading Ability*. Oxford: Oxford University Press.

Readability Calculations, Software, http://www.micropowerandlight.com/rd.html, Micro Power & Light co, downloaded on December 5, 2008.

Rayner, K., & Pollatsek, A. (1994). *The Psychology of Reading*. Englewood Cliffs, NJ: Prentice Hall.

Rosch, E., & Mervis, C. (1975). Family resemblances: Studies in the internal structures of categories. *Cognitive Psychology*, 7, 573-605.

Rosch, E., Mervis, C., Gray, W., Johnson, D., & Boyes-Braem, P. (1976). Basic objects in natural categories. *Cognitive Psychology,* 8, 382-439.

Rosch, E. Human categorization. In N. Warren (Ed.), *Advances in Cross-cultural Psychology (Vol. 1)*. London: Academic Press.

—— (1978). Principles of categorization. In E. Rosch & B. B. Lloyd (Eds.), *Cognition and Categorization,* Social Science Research Council (U.S.).

Rubin, A. (1985). How useful are readability formulas? In J. Osborn, P. T. Wilson, & R. C. Anderson (Eds.), *Reading Education: Foundations for a Literate America*, Lexington, MA: Lexington Books, 61-77.

Schriver, K. A. (2000). Readability formulas in the new millennium: What’s the use? *ACM Journal of Computer Documentation*, 24(3), 138-140.

Thorndike, E. L. (1921). *The Teacher's Word Book*. New York: Teacher’s College Press.

Tversky, B. (1990). Where partonomies and taxonomies meet. In S. L. Tsohatzidis (Ed.), *Meanings and Prototypes: Studies on Linguistic Categorization*. London, 334-344.

Ungerer, F., & Schmid, H. J. (1996). *An Introduction to Cognitive Linguistics*. London/New York, Longman.

Wiener, M., Rubano, M., & Shilkret, R., (1990). A measure of semantic complexity among predications. *Journal of Psycholinguistic Research*, 19(2), 103-123.

WordNet, version 3.0. (2006). Princeton, N.J.: Princeton University. Retrieved from World Wide Web: http://wordnet. princeton.edu/perl/webwn?s=word-you-want.

Appendix A: Online graded readings for American children (Each of 240 words)

Title: Wash Your Hands / Reading level suggested by edHelper: Grades 1-2

Downloaded at http://edhelper.com/ReadingComprehension_29_156.html

It was winter. It was fun in the snow. It was fun playing games. It was not fun being sick. A lot of kids had been sick at school. Some had colds. Some had the flu. Some hurt in the belly. The nurse came to visit each class. She talked about germs. She wanted to stop the germs. She gave tips to keep the germs away. The nurse came to Gabby's class. "Hi!!" she said. "How are you today?" "Fine," the children said. "Have any of you been sick this winter?" the nurse asked. Most of them raised their hands. "Do any of you want to be sick again?" the nurse asked. The kids shook their heads. It is no fun to be sick. "What makes you sick?" the nurse asked. "Germs!" the kids yelled. "That's right!" the nurse said. "What is the best thing you can do to keep germs away?" the nurse asked. The kids did not know what to say. The nurse smiled. "This is an easy one," she said, but the kids were still not sure. "Wash your hands," the nurse told them. "Germs are all over. You spread germs when your hands are dirty. Maybe you touched someone who was sick. Maybe you sneezed or coughed. Maybe you touched a dirty diaper. Maybe some food you touched had germs. If you don't wash the germs away and you touch your eyes or nose or mouth, you may get sick."

Title: The MDA Carnival Package Arrives / Reading level suggested by edHelper: Grades 3–4

Downloaded at http://edhelper.com/ReadingComprehension_29_183.html

Eli ran through the back door. "It's here, it's here," he said, waving a large envelope in the air. "What's here?" asked his sister, Sarah, as she reached into the cabinet to grab a package of cookies. "The carnival kit," said Eli as he sat down at the kitchen counter and started tearing the package apart. Sarah poured herself a glass of milk and put some cookies on a plate. Then she walked over to the counter, placed her snack down, and sat in the chair next to Eli, watching him pull the contents out of the package. "Why do you need a carnival kit?" "I'm going to have a Muscular Dystrophy Association Carnival," said Eli proudly. "It's going to raise lots of money for the MDA, just like Jerry Lewis." "Well la-de-da," said Sarah snippily. "I doubt you'll raise millions of dollars!" Eli glared at his sister. "Maybe not millions, but I'll bet I can raise thousands!" Sarah swallowed a bite of her cookie and washed it down with a gulp of milk. "You have big dreams!" She shook her head. Then she grabbed another cookie and popped it into her mouth. As Sarah continued to eat her snack, Mom came in the back door with an armload of groceries. "Hi, Mom," said Eli. "Look what came today!" "The carnival package!" said Mom, placing the bags of groceries on the counter. "That's great. Have you looked it over yet?"

Title: Physical Therapists / Reading level suggested by edHelper: Grades 4–6

Downloaded at http://edhelper.com/ReadingComprehension_29_192.html

Marilyn was a great-grandmother. One day she took a terrible fall. She needed surgery for her broken hip. She went to a rehabilitation hospital to heal. Who helped her to walk again? A physical therapist did. Joseph was walking to his car. He slipped on a patch of ice. He broke his leg. He needed to use crutches for a bit. Who helped him learn to use them? A physical therapist did. Nicholas is a year old. He was born with a disorder that has hindered his motor development. He attends a special gym three days a week where he can strengthen his muscles by playing with balls, benches, swings, and slides. Who helps him play and grow strong? A physical therapist does. So what is physical therapy? It is a special medical treatment that helps individuals move their bodies. Who does it help? It is meant for those with disabilities, illnesses, or injuries that interfere with movement. It is meant to keep the effects of a disability to a minimum, to help someone feel better, and to speed up recovery. Physical therapy is prescribed by doctors in the orthopedic, neurological, heart, and respiratory fields. Physical therapy has been around for a long time. Ancient people believed in parts of it. It became popular in the United States after the outbreak of World War I. The first school of physical therapy was at the Walter Reed Army Hospital in Washington,

Title: Gift of Horses / Reading level suggested by edHelper: Grades 7-8

Downloaded at http://edhelper.com/ReadingComprehension_29_180.html

When you hear the word “hippotherapy” what do you think of? A hippo in a hot tub? Perhaps a hippo getting a massage? It’s nothing like that at all. Hippo is the Greek word for horse. So hippotherapy is the use of horses to improve health. You may not have heard of it before, but using horses as part of physical therapy began in the U.S. and Germany late in the 1940s. It was slow to catch on, but now there are more than 600 centers for equine or hippotherapy in the United States alone. The benefits of riding horses have been known since the fifth century B.C. Those who rode horses often had better balance, muscular strength, and confidence. In spite of that, the use of horses as part of a regular therapy program has been around less than 60 years, and most of those for less than 20. What is it about riding horses that helps people? The movement of a horse’s hips as it walks is very similar to that of a person’s walk. The gentle movement you feel as you ride helps to exercise many of the same muscles used by humans, but without the effort. For those who cannot walk or who have difficulty walking, those muscles can become weak or atrophied from lack of use. Taking part in a riding therapy program works because those muscles are gently exercised. This builds strength in the

Title: Americans with Disabilities Act in 1990 / Reading level suggested by edHelper: Grades 9-12

Downloaded at http://edhelper.com/ReadingComprehension_29_182.html

America is called the "Land of Opportunity." The Statue of Liberty invites many to our country with the words: "Give me your tired, your poor, Your huddled masses yearning to breathe free, The wretched refuse of your teeming shore. Send these, the homeless, tempest-tossed to me. I lift my lamp beside the golden door." For many groups of people the battle for liberty, civil rights, and equal access has been hard fought. This includes people who are disabled. After World War I, many soldiers returned from battle permanently disabled. Prior to this war, the government granted a pension to disabled veterans. However, the help they needed to readjust to life with a disability was missing. The government stepped in to help. Now disabled veterans were given the opportunity to learn skills needed to find work and regain their daily activities. However, disabled non-veterans were still without assistance until 1935. Under the direction of Franklin D. Roosevelt (who himself was disabled due to polio), the Social Security Program was formed. This program included payments to the permanently disabled to assist them in living. After centuries of being thought of as "burdens to society," public sentiment towards the disabled began to change. However, change is difficult. The barriers in society for the disabled to overcome were tremendous. Not only were there deep-seated fears and misunderstandings in the minds of people, there were physical barriers that needed to be changed. People who used

Appendix B: Sanmin high school English textbook readings (Each of 397 words)

Title: How Does It Taste? / Reading Level: Book 1 Lesson 2

Does milk taste the same as orange juice? Of course not! Does fish taste like chicken? Not at all. But how do you know? What tells you they are different? Is it your tongue? Maybe you think so. But guess again. We do taste things with our tongues; that's true. But the smell of food has a lot to do with its taste, too. We taste foods with our noses as well as our tongues. In fact, the nose has more to do with taste then the tongue. Scientist say that your tongue can recognize only four tastes. It can tell if something is sour (like vinegar) or bitter (like soap). But that's all. To tell different foods apart, we also have to use our noses. Can you remember a time when you had a bad cold? Your food tasted very plain then. It seemed to have little taste at all. That wasn't because your tongue wasn't working. It was because your nose was stopped up. You couldn't smell the food, and that made it seem tasteless. You can prove this to yourself. Try eating something while you pinch your nose shut. It won't seem to have much taste. Here's another test. It shows how important the nose is in tasting. First you blindfold a person. Then you put a piece of potato in his mouth. You tell him to chew it. At the same time, you hold a piece of apple under his nose. Then ask what food is in his mouth. Most people will say, "An apple." The smell of the apple fools them. The test works best when two foods feel the same in the mouth. It won't work well with apple and orange slices. They don't feel alike. What about the eyes? Do they help us taste? Sometimes they may. The way a food looks can make a difference in its taste. Sometimes we taste what we expect to taste. Here's a test to show that: Get some orange food coloring. Mix some into milk. It does not change the taste. Now ask people to taste the orange milk. Ask if it tastes all right. Many people will say it tastes odd. Because it looks odd, they expect an odd taste. And so it tastes odd to them. So you see, it's not only the tongue that does the tasting!

Title: Losing Our Languages / Reading Level: Book 3 Lesson 2

The time may soon come when we say goodbye to most of the world's languages. Today humans express themselves in over 6,000 different languages, but that is quickly changing. Many experts predict that over half of these languages will disappear within the next 50 years. After 100 years, the world may use only a dozen major languages. Why? When people from different cultures live and work together much more than before, change takes place. The languages of the world's dominant cultures are replacing the languages of the smaller cultures. You're learning English right now. Could this be the beginning of the end for the Chinese language? Of course not. *Mandarin* remains the healthy, growing language at the heart of Chinese culture. Mandarin steadily continues to spread among Chinese people worldwide. Elsewhere, *Swahili* grows in Africa. Spanish continues to thrive in *South America*. *Hindi* rules

India. And of course almost everyone these days wants to learn English. However, many less common regional languages haven't been so lucky, because most young people have stopped learning them. When less common languages disappear, two factors are to blame: trade and technology. Most international trade takes place in major world languages such as English or Mandarin. Cultures that isolate themselves from international business and major world languages have difficulty prospering. Most children respect their own culture and traditions. But when it comes to getting a job, knowing a major world language if often essential. It may mean the difference between success and failure. For many, using a less common regional language simply isn't very helpful in today's world. Technology affects languages in an even more fascinating way. Modern media such as radio and television give young people in developing countries much knowledge about the world. These young people can learn about places they've never visited. Their minds open to new events and ideas. This knowledge doesn't come in words from the mouths of their parents or the elders in their community. It usually comes in the language of a dominant culture. It's not surprising then that young people are drawn away from their regional languages. Many benefits come when different cultures begin to share a common language. Instead of struggling for words, people can quickly share ideas and work together. Knowing the same language gives people from different places common ground. A shared language means easier communication and a foundation for trust.

Title: The News / Reading Level: Book 5 Lesson 1

News is an account of events that interest and concern the public. Community residents want to know about a proposed new park in town. The whole nation cares about the devastating earthquake in central Taiwan or an approaching typhoon from the Philippine Sea. To you, information about your friend's flu is news. However, not every story is newsworthy. What is news worthy in one medium may be otherwise in another. The arrival of a new teacher may be reported in the school paper but not in a national newspaper. A hotel fire may make the headlines in local newspapers but not on CNN. What makes a story newsworthy? The question may be answered with the following news elements. Unusualness: A reporter at NBC put it this way: "If an airplane departs on time, it isn't news. If it crashes, regrettably it is." In a nutshell, that comment explains news. News is the different, the unusual, and the out-of-the-ordinary. People sometimes ask, "Why is the news always bad?" Actually, most of the news media include good news, but unusual is more often found in bad news. Significance: Important events, those that affect many people, are news. Some examples are taxes, elections, wars, scientific discoveries, the economy, which are significant in people's lives. Timeliness: Old news isn't news; it's history. People want to hear about the flood while it's happening, not next month when everything has dried out. Proximity: People want to know about nearby events: burglaries in the neighborhood, the proposed regional highway, or the new income tax law. Prominence: When well-known people, buildings, or places are involved, that is news. If you are arrested for shoplifting, it might not make even the local news. But if a movie star is arrested, that's news. Human interest: Stories about ordinary people or animals, humorous or dramatic stories, heartwarming or sad stories often appear in the news because they have human interest: an emotional and personal appeal that draws our attention. Here are two examples that can help you better understand the above concept. United States under Attack Sep 12, 2001 The United States was under attack Tuesday morning, with widespread destruction throughout the East Coast that included at least four commercial jet crashes into significant buildings. The first wave of the attack centered on the World Trade Center in *Manhattan* when a hijacked commercial airline slammed into the second

Appendix C: Lexical relations in the graded readings for American children

Target word	winter	snow	game	kid	school	cold	flu	belly	nurse	class	germ	tip	child	hand	head
Index of synset	0	1	2	0	0	0	0	0	0	1	2	2	0	0	0
Target word length	6	4	4	3	6	4	3	5	5	5	4	3	5	4	4
Number of hyponyms	N/A	2	19	53	98	1	4	2	20	3	N/A	N/A	53	9	8
Number of compounds	N/A	1	6	1	62	1	2	1	13	1	N/A	N/A	6	2	2
Hyponymous length	N/A	6	11.3	8.77	12.9	8	11.5	9.5	10.9	11.7	N/A	N/A	8.77	7	6.13
Length difference	N/A	**2**	**7.32**	5.77	**6.86**	**4**	**8.5**	**4.5**	**5.9**	**6.67**	N/A	N/A	3.77	**3**	**2.13**
Compound ratio	N/A	**50**	**31.6**	1.89	**63.3**	**100**	**50**	**50**	**65**	**33.3**	N/A	N/A	11.3	**22.2**	**25**

Wash Your Hands (Grades 1-2)

thing	diaper	food	eye	nose	mouth	door	envelope	air	sister	cabinet	package	cookie	carnival	kit	kitchen	counter
1	0	0	0	0	1	0	0	1	0	0	0	0	0	0	0	3
5	6	4	3	4	5	4	8	3	6	7	7	6	8	3	7	7
N/A	N/A	689	11	20	8	23	1	N/A	5	5	1	39	2	3	7	4
N/A	N/A	25	3	4	0	18	1	N/A	5	2	0	15	0	2	1	0
N/A	N/A	8.8	7.91	6.3	4.63	9.83	14	N/A	10.4	11.6	4	10.8	9.5	8.67	8	7.75
N/A	N/A	4.8	**4.91**	**2.3**	-0.4	**5.83**	**6**	N/A	**4.4**	**4.6**	-3	**4.77**	1.5	**5.67**	1	0.75
N/A	N/A	3.63	**27.3**	**20**	0	**78.3**	**100**	N/A	**100**	**40**	0	**38.5**	0	**66.7**	14.3	0

The MDA Carnival Package Arrives (Grades: 3-4)

glass	milk	plate	snack	chair	content	money	million	dollar	thousand	bite	gulp	dream	head	mouth	armload	grocery
1	0	3	0	0	0	0	1	0	0	1	0	2	0	0	not	1
5	4	5	5	5	7	5	7	6	8	4	4	5	4	5	found	7
21	24	7	4	48	N/A	48	N/A	25	1	11	N/A	3	8	6	N/A	2
9	24	6	0	31	N/A	1	N/A	24	0	0	N/A	0	2	0	N/A	1
8.71	11.2	10.1	8.5	10.6	N/A	8.33	N/A	13.9	9	5.18	N/A	11	6.13	4.17	N/A	12.5
3.71	**7.17**	**5.14**	3.5	**5.65**	N/A	3.33	N/A	**7.92**	1	1.18	N/A	6	**2.13**	-0.8	N/A	**5.5**
42.9	**100**	**85.7**	0	**64.6**	N/A	2.08	N/A	**96**	0	0	N/A	0	**25**	0	N/A	**50**

bag	grandmother	**fall**	surgery	hip	rehabilitation	hospital	therapist	**car**	patch	ice	**leg**	crutch	bit	disorder	motor	development
0	0	1	3	1	3	0	0	0	0	0	0	0	3	0	1	2
3	11	4	7	3	14	8	9	3	5	3	3	6	3	8	5	11
91	1	2	193	N/A	N/A	32	22	75	25	7	15	N/A	N/A	302	N/A	57
47	0	1	23	N/A	N/A	6	5	20	0	1	4	N/A	N/A	39	N/A	2
8.04	3	7.5	12.9	N/A	N/A	10.6	10.9	8.33	8.36	6.43	8.33	N/A	N/A	14.1	N/A	11.6
5.04	-8	**3.5**	5.9	N/A	N/A	2.59	**1.91**	**5.33**	3.36	3.43	**5.33**	N/A	N/A	6.12	N/A	0.56
51.6	0	**50**	11.9	N/A	N/A	18.8	**22.7**	**26.7**	0	14.3	**26.7**	N/A	N/A	12.9	N/A	3.51

Physical Therapists (Grades 4-6)

gym	**day**	**week**	**muscle**	**ball**	**bench**	swing	swing	**therapy**	treatment	individual	**body**	disability	illness	injury	movement	**effect**
0	0	0	0	0	0	1	3	0	0	0	0	0	0	0	1	0
3	3	4	6	4	5	5	5	7	9	10	4	10	7	6	8	6
N/A	16	8	196	44	10	1	N/A	93	174	2045	27	153	579	124	331	44
N/A	5	2	59	34	3	0	N/A	44	4	1	13	1	2	2	0	10
N/A	7.75	7.5	15.6	9.3	8.2	7	N/A	14.8	13.3	8.45	8.07	12.4	12.9	10.5	7.64	9.86
N/A	**4.75**	**3.5**	**9.58**	**5.3**	**3.2**	2	N/A	**7.83**	4.26	-1.5	**4.07**	2.37	5.86	4.48	-0.4	**3.86**
N/A	**31.3**	**25**	**30.1**	**77.3**	**30**	0	N/A	**47.3**	2.3	0.05	**48.1**	0.65	0.35	1.61	0	**22.7**

minimum	recovery	doctor	field	people	part	outbreak	world	**war**	**school**	word	hippotherapy	hippo	**tub**	**massage**	nothing	**horse**
0	1	0	3	0	0	0	1	0	0	0	Not	1	0	0	0	0
7	8	6	5	6	4	8	5	3	6	4	found	5	3	7	7	5
1	2	99	320	187	1076	3	3	20	98	185	N/A	N/A	4	8	5	134
0	0	11	0	19	1	0	0	5	62	34	N/A	N/A	1	3	0	42
8	5	11.5	12.4	8.84	8.96	9.67	8.33	11.8	12.9	10.5	N/A	N/A	7.25	11	9.2	8.94
1	-3	5.52	7.39	2.84	4.96	1.67	3.33	**8.75**	**6.86**	6.55	N/A	N/A	**4.25**	**4**	2.2	**3.94**
0	0	11.1	0	10.2	0.09	0	0	**25**	**63.3**	18.4	N/A	N/A	**25**	**37.5**	0	**31.3**

Gift of Horses (Grades 7-8)

use	health	part	**therapy**	center	equine	benefit	**century**	**balance**	strength	confidence	program	people	movement	hip	walk	**muscle**
0	0	0	0	2	0	1	0	0	0	0	0	0	1	0	0	0
3	6	4	7	6	6	7	7	7	8	10	7	6	8	3	4	6
65	N/A	1076	93	13	169	5	2	3	40	N/A	153	187	331	N/A	40	196
8	N/A	1	44	8	0	0	1	2	0	N/A	3	17	0	N/A	1	59
13.2	N/A	8.96	14.8	13	8.95	6.6	14	13.3	9.05	N/A	10.4	8.84	7.64	N/A	7.6	15.6
10.2	N/A	4.96	**7.83**	**7**	2.95	-0.4	**7**	**6.33**	1.05	N/A	3.42	2.84	-0.4	N/A	3.6	**9.58**
12.3	N/A	0.09	**47.3**	**61.5**	0	0	**50**	**66.7**	0	N/A	1.96	9.09	0	N/A	2.5	**30.1**

human	effort	difficulty	lack	land	opportunity	statue	liberty	country	word	mass	refuse	shore	**lamp**	**door**	group	people
0	1	0	0	0	0	0	1	0	1	5	0	0	0	2	0	0
5	6	10	4	4	11	6	7	7	4	4	6	5	4	4	5	6
27	103	2	12	8.95	40	9	8	23	N/A	12	N/A	15	68	1	723	187
0	1	0	0	31	0	1	0	3	N/A	0	N/A	2	27	1	19	19
10.6	10.1	7	9.25	12.9	8.85	8.33	7.75	10.3	N/A	8.58	N/A	7.73	10.3	8	8.76	8.84
5.56	4.15	-3	5.25	13.5	-2.2	2.33	0.75	3.35	N/A	4.58	N/A	2.73	**6.32**	**4**	3.76	2.84
0	0.97	0	0	4	0	11.1	0	13	N/A	0	N/A	13.3	**39.7**	**100**	2.63	10.2

Americans with Disabilities Act in 1990 (Grades: 9–12)

battle	**right**	access	world	**war**	soldier	**government**	veteran	skill	work	activity	assistance	direction	polio	**program**	payment	**century**
1	0	1	1	0	0	0	1	0	2	0	0	6	0	1	1	0
6	5	6	5	3	7	10	7	5	4	8	10	9	5	7	7	7
5	109	N/A	3	20	71	36	1	24	22	993	89	15	N/A	18	21	2
0	34	N/A	0	5	2	9	0	0	3	5	0	0	N/A	13	4	1
5.6	13.7	N/A	8.33	11.8	8.79	12.6	11	10.4	9.05	9.62	9.55	12.8	N/A	15.4	10.4	14
-0.4	**8.74**	N/A	3.33	**8.75**	1.79	**2.58**	4	5.38	5.05	1.62	-0.4	3.8	N/A	**8.44**	3.43	**7**
0	**31.2**	N/A	0	**25**	2.82	**25**	0	0	13.6	0.5	0	0	N/A	**72.2**	19	**50**

burden	society	sentiment	change	barrier	fear	misunderstanding	mind
0	0	1	3	1	0	1	2
6	7	9	6	7	4	16	4
4	35	18	2	9	49	N/A	4
0	3	0	0	2	0	N/A	0
7.5	15.8	10.8	14	11.1	8.35	N/A	11.8
1.5	8.8	1.83	8	**4.11**	4.35	N/A	7.75
0	8.57	0	0	**22.2**	0	N/A	0

Appendix D: Lexical relations in the selected readings of Taiwanese high school English textbooks

Target word	**milk**	**orange**	**juice**	fish	chicken	tongue	thing	smell	food	taste	**nose**	scientist	**vinegar**	**soap**	time
Index of synset	0	0	0	1	0	0	7	0	0	0	0	0	0	0	6
Target word length	4	6	5	4	7	6	5	5	4	5	4	9	7	4	4
Number of hyponyms	24	8	8	50	10	N/A	11	18	689	29	20	164	3	14	64
Number of compounds	24	8	6	2	0	N/A	0	0	25	0	4	13	3	12	10
Hyponymous length	11.2	11.8	9.13	9.1	7.7	N/A	8.45	5.94	8.8	7.21	6.3	12.4	11.7	10.1	8.11
Length difference	**7.17**	**5.75**	**4.13**	5.1	0.7	N/A	3.45	0.94	4.8	2.21	**2.3**	3.37	**4.67**	**6.07**	4.11
Compound ratio	**100**	**100**	**75**	4	0	N/A	0	0	3.63	0	**20**	7.93	**100**	**85.7**	15.6

How Does It Taste? (Book 1 Lesson 2)

cold	**test**	person	piece	**potato**	mouth	apple	people	slice	**eye**	way	difference	coloring	time	goodbye	**language**	human
0	1	0	7	0	0	0	0	1	0	0	0	0	2	0	0	0
4	4	6	5	6	5	5	6	5	3	3	10	8	4	7	8	5
1	26	2045	8	10	6	29	187	8	11	38	43	N/A	44	N/A	342	27
1	8	75	0	4	0	5	19	0	3	0	0	N/A	0	N/A	91	5
8	16.5	8.45	7.25	11.6	4.17	10.5	8.84	7.25	7.91	9.34	10.4	N/A	7.93	N/A	11.2	10.6
4	**12.5**	2.45	2.25	**5.6**	-0.8	5.48	2.84	2.25	**4.91**	6.34	0.37	N/A	3.93	N/A	**3.22**	5.56
100	**30.8**	3.67	0	**40**	0	17.2	10.2	0	**27.3**	0	0	N/A	0	N/A	**26.6**	18.5

Losing Our

expert	**year**	world	dozen	people	**culture**	change	beginning	end	heart	**day**	**factor**	**trade**	**technology**	business	difficulty	child
0	0	1	0	0	0	2	0	5	4	1	0	0	0	1	0	0
6	4	5	5	6	7	6	9	3	5	3	6	5	10	8	10	5
260	22	3	1	187	24	764	84	N/A	6	14	10	5	18	160	2	53
5	13	0	0	19	11	10	0	N/A	0	4	3	3	6	6	0	6
10.3	9.09	8.33	7	8.84	17.1	9.2	10.1	N/A	8.8	11	12	13	16	11	7	8.8
4.32	**5.09**	3.33	2	2.84	**10.1**	3.2	1.06	N/A	3.8	**7.8**	**6.4**	**7.6**	**6.1**	2.8	-3	3.8
1.92	**59.1**	0	0	10.2	**45.8**	1.31	0	N/A	0	**29**	**30**	**60**	**33**	3.8	0	11

Languages (Book 3 Lesson 2)

tradition	job	difference	success	failure	way	medium	radio	television	country	knowledge	place	mind	event	idea	word	parent
0	0	0	2	2	0	0	0	0	0	0	0	0	0	0	1	0
9	3	10	7	7	3	6	5	10	7	9	5	4	5	4	4	6
N/A	253	43	2	4	38	102	N/A	4	23	749	93	7	1085	432	N/A	40
N/A	1	0	0	2	0	4	N/A	2	3	9	8	2	9	3	N/A	4
N/A	10	10	7	10	9.3	10	N/A	12	10	9.7	8.62	10.4	9.2	10	N/A	7.1
N/A	7.3	0.4	0	**3**	6.3	4.4	N/A	**2.3**	3.3	0.7	3.62	**6.43**	4.2	6	N/A	1.1
N/A	0.4	0	0	**50**	0	3.9	N/A	50	13	1.2	8.6	**28.6**	0.8	0.7	N/A	10

mouth	elder	community	benefit	ground	communication	foundation	trust	news	account	event	community	resident	park	town	nation	earthquake
1	0	0	1	3	0	0	2	1	0	0	0	0	0	0	1	0
5	5	9	7	6	13	10	5	4	7	5	9	8	4	4	6	10
8	3	23	5	1	171	5	2	30	22	1085	23	18	2	8	17	9
0	0	1	0	1	6	0	0	11	0	9	1	0	2	6	0	1
4.6	5.3	8.3	6.6	12	10	8	11	9	11	9.2	8.3	8.1	11	7.9	8.5	11
-0	0.3	-1	-0	**6**	-3	-2	6	**5**	4.2	4.2	-1	0.1	**7**	**3.9**	2.5	0.7
0	0	4.3	0	**100**	3.5	0	0	**37**	0	0.8	4.3	0	**100**	**75**	0	11

The News (Book 5 Lesson 1)

typhoon	sea	information	friend	flu	story	medium	arrival	teacher	school	newspaper	hotel	fire	headline	question	element	unusualness
0	0	0	0	0	4	0	0	0	0	0	0	0	0	0	0	0
7	3	11	6	3	5	6	7	7	6	9	5	4	8	8	7	11
N/A	1	250	37	4	15	102	3	40	98	7	30	17	9	N/A	17	19
N/A	0	5	3	2	0	4	0	12	62	1	3	12	0	N/A	1	0
N/A	8	9.8	7.4	12	9.6	10	7.3	11	13	7.6	8.2	8.8	9	N/A	12	9.7
N/A	5	-1	1.4	**8.5**	4.6	4.4	0.3	**3.9**	**6.9**	-1	3.2	**4.8**	1	N/A	4.5	-1
N/A	0	2	8.1	**50**	0	3.9	0	**30**	**63**	14	10	**71**	0	N/A	5.9	0

reporter	**airplane**	**time**	**nutshell**	**comment**	**significance**	**example**	**tax**	**election**	**war**	**discovery**	**economy**	**life**	**timeliness**	**history**	**flood**	**month**
0	0	6	0	0	0	0	0	0	0	1	0	1	0	0	0	0
8	8	4	8	7	12	7	3	8	3	9	7	4	10	7	5	5
8	55	64	N/A	35	9	13	64	10	20	N/A	29	N/A	N/A	N/A	8	111
2	2	10	N/A	1	0	0	32	5	5	N/A	8	N/A	N/A	N/A	4	5
12	12	8.1	N/A	7.6	10	8.8	9.8	11	12	N/A	13	N/A	N/A	N/A	10.5	7.02
3.6	3.6	4.1	N/A	0.6	-2	1.8	**6.8**	**3.3**	**8.8**	N/A	**5.9**	N/A	N/A	N/A	**5.5**	2.02
25	3.6	16	N/A	2.9	0	0	**50**	**50**	**25**	N/A	**28**	N/A	N/A	N/A	**50**	4.5

proximity	**burglary**	**neighborhood**	**highway**	**income**	**law**	**prominence**	**building**	**place**	**shoplifting**	**movie**	**star**	**interest**	**animal**	**appeal**	**attention**	**concept**
0	0	0	0	0	1	0	0	0	0	0	1	0	0	1	2	0
9	8	12	7	6	3	10	8	5	11	5	4	8	6	6	9	7
N/A	4	9	23	94	19	11	351	93	N/A	37	1	19	393	3	2	380
N/A	0	0	4	9	11	0	10	8	N/A	2	1	0	11	0	0	4
N/A	13	7.22	9.43	10	12.9	9.27	9.5	8.6	N/A	9.7	9	9.2	8.5	9.7	6	12
N/A	5	-4.8	2.43	4.04	**9.89**	-0.7	1.5	3.6	N/A	4.7	**5**	1.2	2.5	3.7	-3	4.9
N/A	0	0	17.4	9.57	**57.9**	0	2.8	8.6	N/A	5.4	**100**	0	2.8	0	0	1.1

state	**attack**	**morning**	**destruction**	**coast**	**jet**	**crash**	**wave**	**world**	**trade**	**center**	**airline**
0	0	0	1	0	0	1	1	1	0	2	1
5	6	7	11	5	3	5	4	5	5	6	7
9	105	N/A	17	8	8	1	N/A	3	5	13	N/A
4	8	N/A	0	0	7	0	N/A	0	3	8	N/A
13	11.4	N/A	9.12	7.88	7.5	5	N/A	8.33	12.6	13	N/A
8.1	5.37	N/A	-1.9	2.88	**4.5**	0	N/A	3.33	**7.6**	**7**	N/A
44	7.62	N/A	0	0	**87.5**	0	N/A	0	**60**	**61.5**	N/A