Reading in a Foreign Language April 2010, Volume 22, No. 1
ISSN 1539-0578 pp. 201–221

Using glossaries to increase the lexical coverage of television programs

Stuart Webb
Victoria University of Wellington
New Zealand

Abstract

This study examined the extent to which glossaries may affect the percentage of known words (coverage) in television programs. The transcripts of 51 episodes of 2 television programs (*House* and *Grey's Anatomy*) were analyzed using Range (Heatley, Nation, & Coxhead, 2002) to create glossaries consisting of the low-frequency (less frequent than the 3,000 word level) word families that were encountered 10 or more times in each program. The results showed that coverage of the glossaries was 1.31% for *Grey's Anatomy* and 2.26% for *House*. This was greater than coverage of the 3,001–4,000 most frequent word families in both programs. The cumulative coverage including the glossaries at the 3,000 word level increased to 96.00% for *House* and 97.20% for *Grey's Anatomy*. The findings indicate that glossaries have the potential to improve comprehension of television programs.

Keywords: glossary**,** television, vocabulary, coverage, comprehension

Television can be a valuable source of second language (L2) aural input. Research has shown that watching L2 television programs may facilitate incidental vocabulary learning (d'Ydewalle & Pavakanun, 1997; d'Ydewalle & Van de Poel, 1999; Koolstra & Beentjes, 1999; Neuman & Koskinen, 1992; Pavakanun & d'Ydewalle, 1992). L2 learners may also be highly motivated to watch television and movies for language learning (Chapple & Curtis, 2000). This is supported by a study of European and Chinese language learners. Gieve and Clark (2005) found that watching television was the second most commonly used self-directed learning strategy among European learners and the fourth most commonly used strategy among Chinese learners. If learners were to watch television regularly, it may have a significant effect on vocabulary size (Webb & Rodgers, 2009a) because research has shown that increasing the number of times words are encountered in context increases the potential for vocabulary learning (Horst, Cobb, & Meara, 1998; Jenkins, Stein, & Wysocki, 1984; Rott, 1999; Saragi, Nation, & Meister, 1978; Waring & Takaki, 2003; Webb, 2007).

Despite its value as a resource for language learning, there is relatively little research investigating the relationship between L2 learning and television. This may be because L2 television programs are considered too difficult for many learners to understand. Webb and Rodgers (2009a) suggest that a vocabulary size of the most frequent 3,000 word families is necessary for comprehension of L2 television programs. This would indicate that learning with

L2 television may be suitable for many L2 learners. However, the speed of the dialogue, the pronunciation of words which have previously only been encountered in text, and the amount of L2 aural input may still make comprehension of L2 television programs difficult for learners with the appropriate vocabulary size. One way in which comprehension might be increased is through the use of glossaries.

The aim of the present study is to investigate the potential that glossaries may have to improve coverage of L2 television programs. The transcripts of 51 episodes of two television programs were analyzed to determine the most frequent word families that might be unknown to L2 viewers. The coverage that the glossaries represent and the coverage of the programs at the 3,000 word level were examined to determine the extent to which glossaries may increase coverage. Determining whether glossaries have a significant effect on comprehension of television programs experimentally in a controlled treatment was beyond the scope of this study. However, research on the effects of coverage on written (Hu & Nation, 2000; Laufer, 1989) and aural texts (Bonk, 2000) indicates that increasing coverage above 90% is likely to increase comprehension. If glossaries represent a significant amount of coverage then they could have a positive effect on comprehension. Glossaries may reduce the lexical demands of television, and increase comprehension and the potential for incidental vocabulary learning.

Background

Glossing is when text is enhanced by providing the first language (L1) or L2 meanings of difficult words within the text. Glosses are typically found in the margins of the text or at the end of the text in a glossary. Glossaries are commonly found in books such as graded readers which are designed for language learners. In his analysis of the Longman Bridge Series (texts adapted for language learners), Nation (2001) reports that the glossaries at the back of the books ranged from 120 to 600 words, and that the glossed words were not bolded or marked in the texts. The benefit of using unmarked glosses is that they do not disrupt reading and allow learners to simply check words when necessary.

Research investigating the effects of glossing on comprehension has been inconsistent. Several studies have found that glossing has improved comprehension (Davis, 1989; Hulstijn, 1992; Jacobs, 1994; Leffa, 1992; Watanabe, 1997) while two have found no significant effect (Holly & King, 1971; Jacobs, Dufon, & Fong, 1994). One reason for the inconsistent results might be that because there are many factors which affect comprehension, it can be difficult to determine the effects of a single factor (Webb, 2009). Research has also shown that glossing may lead to vocabulary learning (Holley & King, 1971; Hulstijn, Hollander, and Greidanus, 1996; Jacobs, Dufon & Fong, 1994; Watanabe, 1997). In two of those studies (Holley & King, 1971; Jacobs, Dufon & Fong, 1994) significant results were found on immediate posttests but not on delayed tests indicating that gains may be short-lived. Holley and King (1971) found that the placement of the gloss (in the margin or in a glossary) did not influence vocabulary learning and comprehension.

Nation (2001) suggests that glossing is useful for several reasons. First, it may allow more difficult texts to be used by reducing the vocabulary size necessary for adequate comprehension.

Second, by providing the meanings of words which are likely to be unknown, glossing may draw attention to those words and facilitate vocabulary learning. Third, glosses may decrease the time it takes learners to read a text and increase comprehension.

One way to measure the value of a glossary is to examine the coverage that it represents. Coverage is a valuable measure because it may indicate the vocabulary size necessary to understand a text as well as whether incidental vocabulary learning is likely to occur. Laufer and Sim (1985) suggest that coverage may have the greatest effect on whether or not discourse is understood. However, research examining the effects of coverage on comprehension has been inconsistent. Studies have indicated that the coverage necessary for adequate comprehension may range from 90% to 98% depending on the modality (listening or reading) and the type of text. Laufer (1989) found that 95% coverage provided reasonable comprehension of a general academic L2 text. Hu and Nation (2000) found that 98% coverage was sufficient for adequate unassisted L2 reading comprehension of a relatively easy fiction text. This was supported by Carver's (1994) L1 study which indicated that 98%–99% coverage provided adequate L1 comprehension of a text. In contrast, Bonk (2000) found that with adequate coping strategies, L2 learners could have adequate L2 aural comprehension of short texts at far below 95% coverage. Estimates of the amount of coverage necessary for incidental vocabulary learning have also varied. Liu and Nation (1985) suggest that 95% coverage is necessary for learners to successfully guess words from context, and Nation (2001) suggests that 98% coverage is ideal for guessing words from written context.

In a study examining the vocabulary in 88 television programs, Webb and Rodgers (2009a) suggest that 95% coverage may be sufficient for comprehension of television programs. They report that comprehension of television programs may be easier than written text and conversation because the vocabulary heard in television programs is supported by visual input. Research has shown that drawings improved listening comprehension for lower level learners but had no effect for advanced learners (Mueller, 1980), and that learning with video may be more effective than learning with pictures (Hanley, Herron, & Cole, 1995; Secules, Herron, & Tomasello, 1992). In a study examining the effects of visual input on L2 comprehension of television news stories, Gruba (2004) found that there were five ways visual input could facilitate comprehension: helping to identify text type or genre, initiating macrostructure, generating hypotheses, confirming interpretations, and refining interpretations among plausible meanings. He also found that at some points in a television program, visual input can also have no effect or hinder comprehension.

The present study aimed to gain insight into the relative value of glossaries for improving comprehension of television programs. Specifically, the study was designed to investigate the effect of glossaries, which consisted of the low-frequency word families that occurred most often in television programs, on coverage. Meara (1991) took a similar approach in a study of radio broadcasts and suggested that providing word lists derived from word frequency data might be a means of reducing the lexical demands of the broadcasts. Research indicates that glossaries can have a positive effect on comprehension (Davis, 1989; Hulstijn, 1992; Jacobs, 1994; Leffa, 1992; Watanabe, 1997) and vocabulary learning (Holley & King, 1971; Hulstijn, Hollander, & Greidanus, 1996; Jacobs, Dufon, & Fong, 1994; Watanabe, 1997). It is important to look at methods of improving comprehension of L2 television programs because there is a strong

argument for using them for language learning. Research on coverage has shown that increasing coverage above 90% may have a positive effect on comprehension. Because glossaries reduce the lexical demands of text by increasing coverage, they may have the potential to increase comprehension of television programs.

The present study addresses the following questions:

1. What is the coverage of glossaries consisting of the low-frequency word families which occurred 10 or more times in episodes of *House* and *Grey's Anatomy*?

2. What is the coverage of *House* and *Grey's Anatomy* at the 3,000 word level including a glossary consisting of the low-frequency word families which occurred 10 or more times in episodes of each program?

3. What is the difference in coverage between glossaries based on the vocabulary found in that program and the vocabulary found in related programs?

Method

The transcripts of 51 episodes of two English language television programs were analyzed in this study. The two programs, *House* and *Grey's Anatomy*, were chosen because they belonged to the same television genre (i.e., drama) and subgenre (i.e., medical drama), and therefore, might contain a similar vocabulary. Both programs are set in hospitals and are about the lives and relationships of a number of doctors. The episodes made up a complete season of each program; 24 of the episodes consisted of the 2005–2006 season of *House* and the remaining 27 episodes consisted of the 2005–2006 season of *Grey's Anatomy*. All of the episodes were approximately 43 minutes in length.

Words that were not spoken such as stage commands, storyline, and speakers' names were removed from the transcripts. Only words which could be heard when watching the programs were included in the transcripts. Contractions, connected speech, and hyphenated words were changed to conform to spellings used in Nation's (2004a) British National Corpus (BNC) word lists. This accounted for 0.34% of the tokens in the programs. For example, *gonna*, *gotta*, and *wanna* were changed to *going to*, *got to*, and *want to*, respectively. If the spellings were not changed, the words would have been classified as being less frequent than the most frequent 14,000 word-families. However, it is important to note that knowing the spellings which conform to the BNC word lists does not ensure that the original spellings would also be known. Contractions may be one of several factors which may affect comprehension and incidental vocabulary learning. Thus, learners may understand *going to* or *got to* but they might not recognize *gonna* or *gotta*. The small percentage of contractions in the transcripts would suggest that it was unlikely to have a significant influence on comprehension. Another factor that may affect comprehension is the amount of multi-word items. Learners may have the most difficulty with multi-word units when core idioms are encountered (Grant & Bauer, 2004). However, in a study of multi-word items, Grant and Bauer found that there were relatively few core idioms and

that most multi-word units were figuratives, which are less likely to present problems for language learners.

Analysis

The Range program (Heatley, Nation, & Coxhead, 2002) was used to analyze the transcripts. Range is a computer program which lists the words that occur in a text according to their frequency. Nation's (2004a) fourteen 1,000-word lists were used with the Range software to show the number of times each word occurred, and the 1,000 word level (1,000–14,000) at which the words occurred. The lists are based on the frequency and range of occurrence of word families in the BNC. The Range program and the word lists can be downloaded from Paul Nation's website (www.victoria.ac.nz/lals/staff/paul-nation/nation.aspx). Level 6 word families, according to Bauer and Nation's (1993) classification of word families, are used in the lists. Level 6 word families include inflections and over 80 derivational affixes. All word stems were free forms not bound forms. Range classifies items which are not found in the most frequent 14,000 word families as *proper nouns* (List 15), *marginal words* (List 16), and *not in the lists*. Proper nouns which were incorrectly classified as items less frequent than the most frequent 14,000 word families (*not in the lists*) were reclassified and added to the proper nouns list.

Procedure

The transcripts were run through the Range program to determine the cumulative coverage at the 3,000 word level for each program, a random episode of each program, and the 51 episodes combined. The coverage of the proper nouns and marginal words (e.g., *ah, oh, huh*) were included in the cumulative coverage. Nation (2006) suggests that proper nouns and marginal words have a lower learning burden and are more easily learned than typical word families. Webb and Rodgers (2009a, 2009b) took the same approach in their analyses of movies and television programs suggesting that viewers who know the 3,000 most frequent word families should be able to recognize the proper nouns and marginal words.

The Range output was examined to find the most frequent word families from the 4,000 to 14,000 word levels and *not in the lists* (less frequent than the 14,000 word level), the number of times those word families were encountered, and the coverage those word families represented. Glossaries made up of the word families that were encountered 10 or more times in the 4,000 to 14,000 word levels and *not in the lists* were created for each program. A glossary made up of word families that were encountered 20 or more times was also created for the medical genre (all 51 episodes of *House* and *Grey's Anatomy*). The coverage of the glossaries was added to the cumulative coverage of each program and a random episode of each program to determine the potential coverage of the television programs at the 3,000 word level with glossaries.

Results

Grey's Anatomy

The analysis of the transcripts revealed that there were 104 word families in the low-frequency word lists that were encountered 10 or more times in the 27 episodes of *Grey's Anatomy*. The frequency of the items in the program makes them useful to include in a glossary (Nation, 2004b). All of these items are shown in Appendix 1. The glossary primarily consisted of Step 3 and Step 4 technical vocabulary of medicine using Chung and Nation's (2003) rating scale for classifying technical vocabulary. Step 3 items are closely related to the topic (in this case medicine) but are also found in general language. Many of these items related to the body or body parts such as *abdomen, belly, fluid, fracture, hormone, liver, uterus,* and *valve*. Many other items might only be found in the field of medicine making them Step 4 items (words which convey meanings specific to the topic and rarely occur in general language). Examples of Step 4 items are *c-section, CT, enema, morphine, scalpel, surgical, suture, transplant, vitals,* and *x-ray*. About a third of the items did not relate to medicine.

The 104 word families in the glossary accounted for 1.31% of the tokens in *Grey's Anatomy*. The number of encounters with those words ranged from 10, for 16 words, to a maximum of 70 encounters with *intern*. The mean number of encounters with the word families was 17. In a randomly selected episode of *Grey's Anatomy*, 27 of the 104 word families were found in the transcript. The number of encounters with words from the glossary ranged from 1 to 9 in the episode. The glossary accounted for 1.17% of the coverage of the single episode.

Table 1. *Cumulative coverage including proper nouns and marginal words with and without the 104-item glossary for the 27 episodes of* Grey's Anatomy

Word list	Tokens percentage	Coverage (marginal words and proper nouns)	Coverage (marginal words, proper nouns, and glossary)
1,000	86.20	90.10	91.41
2,000	3.89	93.99	95.30*
3,000	1.90	95.89*	97.20
4,000	0.86	96.75	98.06**
5,000	0.60	97.35	98.66
6,000	0.35	97.70	99.01
7,000	0.27	97.97	99.28
8,000	0.19	98.16**	99.47
9,000	0.18	98.34	99.65
10,000	0.14	98.48	99.79
11,000	0.16	98.64	99.95
12,000	0.09	98.73	
13,000	0.08	98.81	
14,000	0.06	98.87	
Proper nouns	2.63		
Marginal words	1.27		
Not in the lists	1.15		
Glossary	1.31		
Tokens	138,561		

Note. *reaching 95% coverage, **reaching 98% coverage.

The percentage of tokens in the fourteen 1,000-word lists and the cumulative coverage with and without the glossary for *Grey's Anatomy* are shown in Table 1. The 27 episodes consisted of 138,561 tokens. The first column shows the percentage of running words found in each of the

1,000 word lists. For example, the most frequent 1,000 word families accounted for 86.20% of the running words, the second 1,000 word families accounted for 3.89%, and the proper nouns and marginal words accounted for 2.63% and 1.27%, respectively. The second column shows the cumulative coverage including proper nouns and marginal words. It is important to note that reaching these coverage points assumes that viewers are able to recognize the proper nouns and understand the marginal words. Table 1 shows that a vocabulary size of the most frequent 3,000 word families plus proper nouns and marginal words is sufficient to reach 95% coverage, and a vocabulary size of the most frequent 8,000 word families plus proper nouns and marginal words is sufficient to reach 98% coverage. The third column of the table reveals that if learners have access to the 104-item glossary, a vocabulary size of the most frequent 2,000 word families plus proper nouns and marginal words is sufficient to reach 95% coverage, and a vocabulary size of the most frequent 4,000 word families plus proper nouns and marginal words is sufficient to reach 98% coverage.

Table 2. *Cumulative coverage including proper nouns and marginal words with and without the 104-item* Grey's Anatomy *glossary for a single episode of the program*

Word list	Tokens percentage	Coverage (marginal words and proper nouns)	Coverage (marginal words, proper nouns, and glossary)
1,000	86.75	89.37	90.54
2,000	3.99	93.36	94.53
3,000	2.50	95.86*	97.03*
4,000	1.06	96.92	98.09**
5,000	0.62	97.54	98.71
6,000	0.10	97.64	98.81
7,000	0.39	98.03**	99.20
8,000	0.12	98.15	99.32
9,000	0.15	98.30	99.47
10,000	0.06	98.36	99.53
11,000	0.15	98.51	99.68
12,000	0.15	98.66	
13,000	0.19	98.85	
14,000	0.02	98.87	
Proper nouns	2.08		
Marginal words	0.54		
Not in the lists	1.14		
Glossary	1.17		
Tokens	5,193		

Note. *reaching 95% coverage, **reaching 98% coverage.

The percentage of tokens in the different word lists and the cumulative coverage with and without the glossary for the single episode of *Grey's Anatomy* are shown in Table 2. The single episode consisted of 5,193 tokens. The most frequent 1,000 word families accounted for 86.75% of the tokens and the glossary accounted for 1.17%. The table shows that without a glossary, a vocabulary size of the most frequent 3,000 word families plus proper nouns and marginal words is sufficient to reach 95.86% coverage, and a vocabulary size of the most frequent 7,000 word families plus proper nouns and marginal words is sufficient to reach 98.03% coverage. With the glossary, a vocabulary size of the most frequent 3,000 word families plus proper nouns and marginal words is sufficient to reach 97.03% coverage, and a vocabulary size of the most

frequent 4,000 word families plus proper nouns and marginal words is sufficient to reach 98.09% coverage.

House

There were 128 word families encountered 10 or more times in the 24 episodes of *House*. These items are shown in Appendix 2. The glossary consisted primarily of Step 3 and Step 4 (Chung & Nation, 2003) technical vocabulary of medicine. Twenty-nine of the word families were also found in the *Grey's Anatomy* glossary indicating that there was significant overlap in vocabulary between the two programs. All of the word families found in both glossaries except for *bullet*, *kiss*, *quit*, and *whoa* could be classified as technical vocabulary from the field of medicine.

The 128 word families in the glossary accounted for 2.26% of the running words in *House*. The number of encounters with those words ranged from 10, for 11 word families, to a maximum of 99 encounters with *tumor*. The mean number of encounters was 22. In a randomly selected episode of *House*, 42 of the 128 word families were found in the transcript. The glossary accounted for 1.59% of the coverage of the episode.

Table 3. *Cumulative coverage including proper nouns and marginal words for the 24 episodes of* House *with and without the 128-item glossary*

Word list	Tokens percentage	Coverage (marginal words and proper nouns)	Coverage (marginal words, proper nouns, and glossary)
1,000	84.70	86.34	88.60
2,000	4.88	91.22	93.48
3,000	2.52	93.74	96.00*
4,000	1.30	95.04*	97.30
5,000	0.87	95.91	98.17**
6,000	0.57	96.48	98.74
7,000	0.36	96.84	99.10
8,000	0.41	97.25	99.51
9,000	0.27	97.52	99.78
10,000	0.26	97.78	
11,000	0.22	97.99	
12,000	0.17	98.17**	
13,000	0.08	98.25	
14,000	0.14	98.39	
Proper nouns	1.20		
Marginal words	0.44		
Not in the lists	1.63		
Glossary	2.26		
Tokens	128,295		

Note. *reaching 95% coverage, **reaching 98% coverage.

The percentage of tokens in the different word lists and the cumulative coverage with and without the glossary for *House* are shown in Table 3. The 24 episodes consisted of 128,295 tokens. The most frequent 1,000 word families accounted for 84.70% of the tokens with the percentage consistently decreasing as the word frequency decreases. Table 3 shows the cumulative coverage with and without the 128-item glossary in columns 2 and 3. Without the

glossary, a vocabulary size of the most frequent 4,000 word families plus proper nouns and marginal words is sufficient to reach 95% coverage, and a vocabulary size of the most frequent 12,000 word families plus proper nouns and marginal words is sufficient to reach 98% coverage. In contrast, the third column of the table shows that if learners use the glossary, a vocabulary size of the most frequent 3,000 word families plus proper nouns and marginal words is sufficient to reach 96% coverage, and a vocabulary size of the most frequent 5,000 word families plus proper nouns and marginal words is sufficient to reach 98% coverage.

Table 4 shows the percentage of tokens in the word lists and the cumulative coverage with and without the 128-item glossary for the single episode of *House*. There were 5,272 tokens in the episode. The glossary accounted for 1.59% of the running words in the episode. Without the glossary, a vocabulary size of the most frequent 5,000 word families plus proper nouns and marginal words was necessary to reach 95% coverage, and a vocabulary size of the most frequent 13,000 word families plus proper nouns and marginal words is sufficient to reach 98% coverage. In contrast, if learners have access to the glossary, a vocabulary size of the most frequent 3,000 word families plus proper nouns and marginal words is sufficient to reach 95% coverage, and a vocabulary size of the most frequent 7,000 word families plus proper nouns and marginal words is sufficient to reach 98% coverage.

Table 4. *Cumulative coverage including proper nouns and marginal words of a single episode of* House *with and without the 128-item glossary*

Word list	Tokens percentage	Coverage (marginal words and proper nouns)	Coverage (marginal words, proper nouns, and glossary)
1,000	84.77	86.22	87.81
2,000	5.25	91.47	93.06
3,000	2.33	93.80	95.39*
4,000	0.97	94.77	96.36
5,000	0.70	95.47*	97.06
6,000	0.57	96.04	97.63
7,000	0.40	96.44	98.03**
8,000	0.64	97.08	98.67
9,000	0.44	97.52	99.11
10,000	0.11	97.63	99.22
11,000	0.19	97.82	99.41
12,000	0.13	97.95	
13,000	0.08	98.03**	
14,000	0.09	98.12	
Proper nouns	1.11		
Marginal words	0.34		
Not in the lists	1.88		
Glossary	1.59		
Tokens	5,272		

Note. *reaching 95% coverage, **reaching 98% coverage.

Medical Genre

There were 93 word families encountered 20 or more times in the low-frequency word lists (4,000–14,000 plus *not in the lists*), 153 word families encountered 15 or more times, and 272

word families encountered 10 or more times in the 51 episodes of *House* and *Grey's Anatomy*. Because the 93-item glossary was most similar in size to the individual glossaries of *House* and *Grey's Anatomy*, and it would be the most manageable size for learners, it was created for comparison. It is shown in Appendix C. The medical genre glossary had 54 of the same word families as the *Grey's Anatomy* glossary and 68 of the same items as the *House* glossary. The number of encounters with the word families ranged from 20 encounters with 7 items to a maximum of 151 encounters with *tumor*. The mean number of encounters was 37. There were 3,461 encounters with words from the 93-item genre glossary in the 51 episodes.

The percentage of tokens in the word lists and the cumulative coverage with and without the 93-item glossary for the 51 episodes of *House* and *Grey's Anatomy* are shown in Table 5. The glossary accounted for 1.30% coverage. Without a glossary, a vocabulary size of 4,000 word families plus proper nouns and marginal words is sufficient to reach 95% coverage and a vocabulary size of 10,000 words plus proper nouns and marginal words is sufficient to reach 98% coverage. With the glossary the vocabulary size necessary to reach 95% decreases to 3,000 word families plus proper nouns and marginal words, and the vocabulary size necessary to reach 98% coverage decreases to 6,000 word families and proper nouns and marginal words.

Table 5. *Cumulative coverage including proper nouns and marginal words of the medical genre with and without the 93-item glossary*

Word list	Tokens percentage	Coverage (marginal words and proper nouns)	Coverage (marginal words, proper nouns, and glossary)
1,000	85.48	88.29	89.59
2,000	4.36	92.65	93.95
3,000	2.20	94.85	96.15*
4,000	1.07	95.92*	97.22
5,000	0.73	96.65	97.95
6,000	0.45	97.10	98.40**
7,000	0.32	97.42	98.72
8,000	0.30	97.72	99.02
9,000	0.22	97.94	99.24
10,000	0.20	98.14**	99.44
11,000	0.19	98.33	99.63
12,000	0.12	98.45	99.75
13,000	0.08	98.53	99.83
14,000	0.10	98.63	99.93
Proper nouns	1.94		
Marginal words	0.87		
Not in the lists	1.38		
Glossary	1.30		
Tokens	266,856		

Note. *reaching 95% coverage, **reaching 98% coverage.

Table 6 contrasts the cumulative coverage of *Grey's Anatomy* with the glossary derived exclusively from the episodes of the program and with the medical genre glossary. There were 1,410 encounters with words from the 93-item genre glossary in the 27 episodes of *Grey's Anatomy*. This represented 1.02% of the tokens in *Grey's Anatomy*. This is 0.29% less than coverage of the glossary derived from the 27 episodes. Eighty-six of the 93 items in the genre

glossary were found in *Grey's Anatomy* with the number of encounters ranging from 1 to 70. The mean number of encounters was 15. A vocabulary size of 2,000 word families provided 95% coverage with each glossary. However, a vocabulary size of 4,000 word families provided 98% coverage with the program glossary, and a vocabulary size of 5,000 word families is necessary to reach 98% coverage using the genre glossary.

In the single episode of *Grey's Anatomy*, 27 of the 93 items in the genre glossary were encountered a total of 45 times. This represented 0.87% of the tokens in the program, which is 0.30% less than the *Grey's Anatomy* glossary. A vocabulary size of 3,000 word families provided 95% coverage with each glossary. However, a vocabulary size of 4,000 word families provided 98% coverage with the program glossary, and a vocabulary size of 5,000 word families is necessary to reach 98% coverage with the genre glossary.

Table 6. *Cumulative coverage including proper nouns and marginal words of* Grey's Anatomy *with the program and genre glossaries*

Word list	*Grey's Anatomy*			Single episode of *Grey's Anatomy*		
	No glossary	Program glossary	Genre glossary	No glossary	Program glossary	Genre glossary
		104 items	93 items		104 items	93 items
1,000	90.10	91.41	91.12	89.37	90.54	90.24
2,000	93.99	95.30*	95.01*	93.36	94.53	94.23
3,000	95.89	97.20	96.91	95.86	97.03*	96.73*
4,000	96.75	98.06**	97.77	96.92	98.09**	97.79
5,000	97.35	98.66	98.37**	97.54	98.71	98.41**
6,000	97.70	99.01	98.72	97.64	98.81	98.51
7,000	97.97	99.28	98.99	98.03	99.20	98.90
8,000	98.16	99.47	99.18	98.15	99.32	99.02
9,000	98.34	99.65	99.36	98.30	99.47	99.17
10,000	98.48	99.79	99.50	98.36	99.53	99.23
11,000	98.64	99.95	99.66	98.51	99.68	99.38
12,000	98.73		99.75	98.66		99.53
13,000	98.81		99.83	98.85		99.72
14,000	98.87		99.89	98.87		99.74
Glossary		1.31	1.02		1.17	0.87

Note. *reaching 95% coverage, **reaching 98% coverage.

The cumulative coverage of *House* with the glossary derived exclusively from the episodes of the program and with the medical genre glossary is shown in Table 7. There were 2,066 encounters with words from the 93-item glossary in the 24 episodes of *House*. This represented 1.61% of the tokens in the 24 episodes, which is 0.65% less than the coverage of the glossary derived exclusively from the program. Eighty-nine of the 93 words in the genre glossary were encountered in *House* with the number of encounters ranging from 1 to 99.

In the single episode of *House*, there were 69 encounters with the words from the 93-item genre glossary. This represented 1.31% of the tokens in the episode. This is 0.28% less than coverage of the glossary derived from the 24 episodes of *House*. Thirty-four of the 93 items were found in the episode with the number of encounters ranging from 1 to 5. A vocabulary size of 3,000 word families provided 95% coverage with each glossary. However, a vocabulary size of 7,000 word

families provided 98% coverage with the program glossary, and a vocabulary size of 8,000 word families is necessary to reach 98% coverage using the genre glossary.

Table 7. *Cumulative coverage including proper nouns and marginal words of* House *with the program and genre glossaries*

Word list	*House*			Single episode of *House*		
	No glossary	Program glossary 128 items	Genre glossary 93 items	No glossary	Program glossary 128 items	Genre glossary 93 items
1,000	86.34	88.60	87.95	86.22	87.81	87.53
2,000	91.22	93.48	92.83	91.47	93.06	92.78
3,000	93.74	96.00*	95.35*	93.80	95.39*	95.11*
4,000	95.04*	97.30	96.65	94.77	96.36	96.08
5,000	95.91	98.17**	97.52	95.47*	97.06	96.78
6,000	96.48	98.74	98.09**	96.04	97.63	97.35
7,000	96.84	99.10	98.45	96.44	98.03**	97.75
8,000	97.25	99.51	98.86	97.08	98.67	98.39**
9,000	97.52	99.78	99.13	97.52	99.11	98.83
10,000	97.78		99.39	97.63	99.22	98.94
11,000	97.99		99.61	97.82	99.41	99.13
12,000	98.17**		99.78	97.95		99.26
13,000	98.25		99.86	98.03**		99.34
14,000	98.39		100.00	98.12		99.43
Glossary		2.26	1.61		1.59	1.31

Note. *reaching 95% coverage, **reaching 98% coverage.

Discussion

In answer to the first research question, the results showed that coverage of the program glossaries was 1.31% and 2.26% and coverage of the genre glossaries was 1.02% and 1.61% for *Grey's Anatomy* and *House*, respectively. Coverage of the program glossaries was 1.17% and 1.59% for the individual episodes and 0.87% and 1.31% with the genre glossary. Coverage of the genre glossary was 1.30% for the medical genre. The relative value of the glossaries can be seen by comparing the coverage of the glossaries with the coverage of the 1,000 word lists. In the glossary for *Grey's Anatomy*, the 104 word families in the glossary accounted for 1.31% of the running words. This is a greater percentage of words in the episodes than the 3,001 to 4,000 most frequent word families (0.86%). The *Grey's Anatomy* glossary also accounted for greater coverage (1.17%) of the single episode than the fourth 1,000 word list (1.06%). Thus, for both the set of 27 episodes and the single episode, only the first three 1,000 word lists and the list of proper nouns accounted for more running words. The glossary for *House* also accounted for more running words than each of the 4,000–14,000 word lists. The 128-item glossary accounted for 2.26% of the tokens in *House* which was lower coverage than the 3,000 word list (2.52%) but higher coverage than the 4,000 word list (1.30%). In fact the glossary provided greater coverage than the 3,001st to 5,000th word families (2.17%). It also provided 1.59% of the running words in the episode of *House* which was higher coverage than coverage of the 4,000 word list (0.97%) for that episode.

The 93-item genre glossary (1.30%) also provided greater coverage of the genre than the 4,000 word list (1.07%). However, it accounted for slightly lower coverage of the two individual programs. It provided greater coverage of *Grey's Anatomy* (1.02%) than the 4,000 word list (0.86%) and greater coverage of the single episode of *Grey's Anatomy* (0.87%) than the 5,000 word list (0.62%) but lower coverage than the 4,000 word list (1.07%). The genre glossary also provided greater coverage of the 24 episodes of *House* (1.61%) than the 4,000 word list (1.30%), and accounted for higher coverage of the single episode of *House* (1.31%) than the 4,000 word list (0.97%) for that episode.

Taken together, the comparisons between coverage of the glossaries and coverage of the 1,000 word lists provide evidence that glossaries may have great value in assisting comprehension of television programs. The results showed that knowing the low-frequency word families which were encountered most often, may have a greater effect on comprehension than knowing the 3,001st–4,000th most frequent word families in eight of the nine comparisons. The difference in size between the glossaries (104, 128, and 93 items) and the word lists (1,000 items) would suggest that creating and using glossaries based on frequency may be a very effective means of aiding comprehension.

In answer to the second research question, the results indicated that glossaries based on frequency may increase vocabulary coverage to 95% or higher at the 3,000 word level. This is important because as coverage increases from 90% to 95%, learners are likely to have improved comprehension (Hu & Nation, 2000; Laufer, 1989). Webb and Rodgers (2009a) suggest that 95% coverage may be sufficient to understand television programs, and Liu and Nation (1985) report that there is the potential for incidental learning to occur at 95% coverage. The findings showed that coverage of the 24 episodes of *House* increased from a point at which viewers may not have adequate comprehension at the 3,000 word level (93.74%) to a coverage with the program glossary (96.00%) and the genre glossary (95.35%) at which they are more likely to understand. Coverage of the single episode of *House* also increased from 93.80% at the 3,000 word level to 95.39% with the program glossary and 95.11% with the genre glossary. Coverage of the 27 episodes of *Grey's Anatomy* increased from 95.89% at the 3,000 word level to 97.20 with the program glossary and 96.91% with the genre glossary. It should also be noted that coverage of *Grey's Anatomy* increased from 93.99% at the 2,000 word level to above 95% with each of the program and genre glossaries. Coverage of the single episode of *Grey's Anatomy* increased from 95.86% at the 3,000 word level to 97.03% with the program glossary and 96.73% with the genre glossary.

Together, the results indicate that glossaries may have value as tools for aiding comprehension of television programs. Because research indicates that learners are motivated to watch L2 television programs (Gieve & Clark, 2005), and that watching L2 television programs may lead to incidental vocabulary learning (d'Ydewalle & Pavakanun, 1997; d'Ydewalle & Van de Poel, 1999; Koolstra and Beentjes, 1999; Neuman & Koskinen, 1992; Pavakanun & d'Ydewalle, 1992), creating glossaries may be a useful method of making television programs more accessible to L2 learners. The availability of L2 television programs on DVD for rental and purchase in most countries makes them a valuable resource for language learning.

In answer to the third research question, the results showed that the program glossaries provided greater coverage than the genre glossary, but both types of glossaries provided relatively large coverage of the programs. It should be expected that a program glossary provides greater coverage of that program than a genre glossary because items in a program glossary are based on the frequency of occurrence of items in that program. However, the fact that the genre glossary provided greater coverage than the 4,000 word list for both programs suggests that creating genre glossaries may also be effective. The advantage of creating genre glossaries rather than program glossaries is that they reduce workload. The disadvantage is that if the vocabulary in a program included in the genre is not typical of the genre then the glossary is likely to have little value.

Creating and Using Glossaries

There are two ways that glossaries could be developed. One way is to create them based on computer analyses of transcripts. This will provide the most accurate data on the frequency of occurrence of items. Tools created by Paul Nation are particularly useful for this type of analysis. The Range program (Heatley, Nation, & Coxhead, 2002) used together with the fourteen 1,000 word BNC lists (Nation, 2004a) provides a quick and effective analysis of the vocabulary in text. A second approach would be for teachers to watch the programs and create a list of the critical words necessary for comprehension. This would be more time consuming and would not have the same effect on coverage. However, it would take into consideration factors such as storyline and importance of the items for comprehension which computer programs cannot.

Once the items have been selected, teachers must decide what information to include in the glossary. For example, L1 definitions, L2 definitions, pictures, examples, phonetic transcriptions could all be included. Research on glossing indicates that both L1 and L2 glosses can be effective, providing that the meaning of the items is clear (Nation, 2001). The amount of information provided in the glossary should perhaps be dependent on the value of the word for the learners. For example, the vocabulary in *House* and *Grey's Anatomy* is likely to be more technical than in many other programs due to the subject matter. Nation (2001, 2008) suggests that learners should spend the most time on the vocabulary which is most useful to them. These items tend to be the most frequent words which are yet to be learned or words which fill a need such as technical or academic vocabulary. Words in these glossaries such as *enema*, *interferon*, *biopsy*, *tachycardia*, and *benign* may only be partially known to many viewers who watch these programs in their L1 and may have little value to L2 learners unless they aspire to work in the field of medicine. Providing very short easily understood definitions for these items is best. In fact it would be useful when creating glossaries to provide symbols which indicate the frequency or usefulness of the words in the glossary for L2 learners. For example, symbols could indicate that in the 128-item glossary for *House*, words such as *fever*, *liver*, *consent*, and *prescribe* may be most useful as these words are more likely to occur in general language. Teachers could get the frequency information from dictionaries or by running the words through Range using Nation's (2004a) BNC lists. Items which include useful word parts, such as *abnormal*, may also be valuable as they may contribute to learning the meaning of affixes (*ab)* which can help for future learning. Other items such as *amphetamine*, *tachycardia*, and *tuberculosis* are less useful. Minimal definitions such as *a type of medicine or drug* for *amphetamine*, *a problem with the heart* for *tachycardia*, and *a disease* for *tuberculosis* may be sufficient.

Glossaries could also be developed by both teachers and learners. Teachers could provide the target words in a glossary with learners supplying the definitions. Once a glossary has been satisfactorily completed, it can be used again and again and even updated if learners find that there are key words which are not included in the glossary. Over time, a library of programs could be accumulated and graded according to vocabulary load and difficulty. It is important to remember that vocabulary load, while a key factor in comprehension, is not the only factor, and that other factors such as the number of contractions, the speed of the dialogue, the amount of background knowledge, the genre, and the importance of key sentences or phrases in the dialogue may have an influence on comprehension. Andrade (1997) found that enhanced background knowledge of television commercials had a positive effect on comprehension. Providing background information, such as character names and other key proper nouns, and a brief synopsis of programs together with glossaries may be another way to help increase comprehension.

If L2 television viewing is to focus on meaning rather than language, it may be best to guide learners to consult a glossary both before and after viewing. This should keep the focus on meaning and fluency during viewing. Television, like reading, is individual in nature. The use of glossaries fits in well with individual television viewing if learners are able to stop or pause programs and consult a glossary when necessary. In classrooms, glossaries may be less effective because different learners may wish to pause the program and consult their glossary at different times making viewing more intensive then extensive. Glossaries may be most effectively used in self-access centers where DVDs of television programs are available and learners can consult a glossary when needed.

Limitations

The findings in this study indicate that knowing the most frequent 3,000 word families and having access to a glossary based on the frequency of occurrence of items may be sufficient to reach 95% coverage of television programs. Research suggests that the 95% coverage figure may signal that viewers have adequate comprehension and have the potential to learn words incidentally. However, it is important to note that empirical research investigating the effects of coverage of television programs on comprehension and incidental vocabulary learning is lacking. All of the studies which have measured the effects of coverage have examined its effects on comprehension of either written (Carver, 1994; Hu & Nation, 2000; Laufer, 1989) or aural text (Bonk, 2000). Further empirical research needs to look at the effect that coverage has on comprehension and vocabulary learning. In particular, research investigating the coverage necessary to reach different degrees of comprehension of television programs is needed. Determining the vocabulary size necessary to have adequate and more precise comprehension of different programs would provide a more accurate target vocabulary size for learners. It would also be useful to determine the coverage point at which learners are satisfied with viewing L2 programs. If, for example, learners have less precise comprehension of programs at a coverage below 90% but still feel satisfied and motivated to learn with television, does learning occur and if so to what extent? Other questions that would be useful to investigate are as follows:

- Does aural comprehension improve through repeated viewings of different programs?

- Are learners able to incidentally learn vocabulary at a lower coverage?

- What coverage is ideal for incidental vocabulary learning? Nation (2001) suggests that 98% coverage is ideal for reading but more research is needed to look at this question for each of reading, listening, and watching television and movies.

It is also important to note that coverage of television programs is likely to vary from program to program and from episode to episode. Finding one episode that learners can understand does not ensure that they will understand other episodes equally well. Moreover, a glossary may be particularly useful for one episode but less useful for another. Both teachers and learners need to be aware of the variation between programs and episodes. In the initial stages of using television for language learning, teachers should either analyze the vocabulary load of programs with Range or carefully screen programs to ensure that learners who know the most frequent 3,000 word families are likely to have adequate comprehension. Creating glossaries is one method of increasing coverage to a point where adequate comprehension is more likely to occur.

It should also be noted that although vocabulary may be the factor which has the greatest effect on comprehension, it is one of a number of factors which affects comprehension. For example, the link between form and meaning in the imagery, the rate of speech, and the clarity of speech may influence listening comprehension (Rubin, 1994). Research has also shown that there are many factors that may affect reading comprehension such as background knowledge (Stahl, Hare, Sinatra, & Gregory, 1991; Stahl & Jacobson, 1986; Stahl, Jacobson, Davis, & Davis, 1989), the relevance of unknown vocabulary in context (Stahl, 1990), the amount of redundant information (Kameenui, Carnine, & Freschi, 1982), and individual differences (Mezynski, 1983; Stahl, 1990). These may also affect comprehension of television programs. Because of the number of factors which can affect comprehension, it is important to be aware that 100% coverage does not always ensure comprehension.

Acknowledgments

I am most grateful to Paul Nation for his advice, support, and friendship. Paul has not only had a positive influence on all of my research, but he also made doing a PhD and working on research fun. I would also like to thank Michael Rodgers for his work developing the corpus discussed in this study, and the anonymous reviewers for their helpful comments.

References

Andrade, H. S. (1997). The effects of previewing and vocabulary and providing background information on the comprehension of television commercials. *JACET Bulletin*, *28*, 1–16.

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

Bonk, W. (2000). Second language lexical knowledge and listening comprehension. *International Journal of Listening, 14*, 14–31.

Carver, R. P. (1994). Percentage of unknown vocabulary words in text as a function of the relative difficulty of the text: Implications for instruction. *Journal of Reading Behavior, 26*, 413–437.

Chapple, L., & Curtis, A. (2000). Content-based instruction in Hong Kong: Student responses to film. *System, 28*, 419–433.

Chung, T., & Nation, I. S. P. (2003). Technical vocabulary in specialised texts. *Reading in a Foreign Language, 15*, 103–116.

Davis, J. (1989). Facilitating effects of marginal glosses on foreign language reading. *The Modern Language Journal, 73*, 41–48.

d'Ydewalle, G., & Pavakanun, U. (1997). Could enjoying a movie lead to language acquisition? In P. Winterhoff-Spurk & T. Van der Voort (Eds.), *New horizons in media psychology* (pp. 145–155). Opladen, Germany: Westdeutscher Verlag GmbH.

d'Ydewalle, G., & Van de Poel, M. (1999). Incidental foreign-language acquisition by children watching subtitled television programs. *Journal of Psycholinguistic Research, 28*, 227–244.

Gieve, S., & Clark, R. (2005). "The Chinese approach to learning": Cultural trait or situated response? The case of a self-directed learning programme. *System, 33*, 261–276.

Grant, L., & Bauer, L. (2004). Criteria for re-defining idioms: Are we barking up the wrong tree? *Applied Linguistics, 25*, 38–61.

Gruba, P. (2004). Understanding digitized second language videotext. *Computer Assisted Language Learning, 17*, 51–82.

Hanley, J., Herron, C., & Cole, S. (1995). Using video as advance organizer to a written passage in the FLES classroom. *The Modern Language Journal, 79*, 57–66.

Heatley, A., Nation, I. S. P., & Coxhead, A. (2002). Range [Computer software]. Retrieved from http://www.victoria.ac.nz/lals/staff/paul-nation/nation.aspx

Holley, F., & King, J. (1971). Vocabulary glosses in foreign language learning materials. *Language Learning, 21*, 213–19.

Horst, M., Cobb, T., & Meara, P. (1998). Beyond *A Clockwork Orange*: Acquiring second language vocabulary through reading. *Reading in a Foreign Language, 11*, 207–223.

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

Hulstijn, J. (1992). Retention of inferred and given word meanings: Experiments in incidental vocabulary learning. In Arnaud, P. & Béjoint, H., (Eds.), *Vocabulary and applied linguistics* (pp. 113–25). London: Macmillan.

Hulstijn, J. H., Hollander, M., & Greidanus, T. (1996). Incidental vocabulary learning by advanced foreign language students: The influence of marginal glosses, dictionary use, and reoccurrence of unknown words. *The Modern Language Journal, 80*, 327–339.

Jacobs, G. (1994). What lurks in the margin: Use of vocabulary glosses as a strategy in second language reading. *Issues in Applied Linguistics, 5*, 115–37.

Jacobs, G., Dufon, P., & Fong, C. (1994). L1 and L2 glosses in L2 reading passages: Their effectiveness for increasing comprehension and vocabulary knowledge. *Journal of Research in Reading, 17*, 19–28.

Jenkins, J. R., Stein, M. L., & Wysocki, K. (1984). Learning vocabulary through reading. *American Educational Research Journal, 21*, 767–787.

Kameenui, E. J., Carnine, D. C., & Freschi, R. (1982). Effects of text construction and instructional procedures for teaching word meanings on comprehension and recall. *Reading Research Quarterly*, *17*, 367–388.

Koolstra, C., & Beentjes, J. (1999). Children's vocabulary acquisition in a foreign language through watching subtitled television programs at home, *Educational Technology Research and Development*, *47*, 51–60.

Laufer, B. (1989). What percentage of text lexis is essential for comprehension? In C. Lauren & M. Nordman (Eds.), *Special language: From humans thinking to thinking machines*, (pp. 316–323). Clevedon, England: Multilingual Matters.

Laufer, B., & Sim, D. D. (1985). An attempt to measure the threshold of competence for reading comprehension. *Foreign Language Annals, 18*, 405–411.

Leffa, V. J. (1992). Making foreign language texts comprehensible for beginners: An experiment with an electronic glossary. *System*, *20*, 63–73.

Liu Na., & Nation, I. S. P. (1985). Factors affecting guessing vocabulary in context. *RELC Journal*, *16*, 33–42.

Meara, P. M. (1991). *BBC English core curriculum: The lexicon.* London: BBC English.

Mezynski, K. (1983). Issues concerning the acquisition of knowledge: Effects of vocabulary training on reading comprehension. *Review of Educational Research*, *53*, 253–279.

Mueller, G. (1980). Visual contextual cues and listening comprehension: An experiment. *The Modern Language Journal*, *64*, 335–340.

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

Nation, I. S. P. (2004a). A study of the most frequent word families in the British National Corpus. In P. Bogaards & B. Laufer (Eds.), *Vocabulary in a second language: Selection, acquisition, and testing* (pp. 3–13). Amsterdam: John Benjamins.

Nation, I. S. P. (2004b). Vocabulary learning and intensive reading. *EA Journal*, *21*, 20–29.

Nation, I. S. P. (2006). How large a vocabulary is needed for reading and listening? *The Canadian Modern Language Review*, *63*, 59–82.

Nation, I. S. P. (2008). *Teaching vocabulary: Strategies and techniques.* Boston, MA: Heinle.

Neuman, B., & Koskinen, P. (1992). Captioned television as comprehensible input: Effects of incidental word learning from context for language minority students. *Reading Research Quarterly*, *27*, 95–106.

Pavakanun, U., & d'Ydewalle, G. (1992). Watching foreign television programs and language learning. In F. L. Engel, D. G. Bouwhuis, T. Bösser, & G. d'Ydewalle (Eds.), *Cognitive modelling and interactive environments in language learning* (pp. 193–198). Berlin: Springer.

Rott, S. (1999). The effect of exposure frequency on intermediate language learners' incidental vocabulary acquisition through reading. *Studies in Second Language Acquisition*, *21*, 589–619.

Rubin, J. (1994). A review of second language listening comprehension research. *The Modern Language Journal*, *78*, 199–221.

Saragi, T., Nation, I. S. P., & Meister, G. F. (1978). Vocabulary learning and reading. *System*, *6*, 72–78.

Secules, T., Herron, C., & Tomasello, M. (1992). The effect of video context on foreign language learning. *The Modern Language Journal*, *76*, 480–490.

Stahl, S. A. (1990). *Beyond the instrumentalist hypothesis: Some relationships between word meanings and comprehension* [Technical report No. 505]. Urbana: University of Illinois, Center for the Study of Reading.

Stahl, S. A., Hare, V. C., Sinatra, R., & Gregory, J. F. (1991). Defining the role of prior knowledge and vocabulary in reading comprehension: The retiring of number 41. *Journal of Reading Behavior, 23,* 487–508.

Stahl, S. A., & Jacobson, M. G. (1986). Vocabulary difficulty, prior knowledge, and text comprehension. *Journal of Reading Behavior, 18,* 309–323.

Stahl, S. A., Jacobson, M. G., Davis, C. E., & Davis, R. L. (1989). Prior knowledge and difficult vocabulary in the comprehension of unfamiliar text. *Reading Research Quarterly, 24,* 27–43.

Waring, R., & M. Takaki. (2003). At what rate do learners learn and retain new vocabulary from reading a graded reader? *Reading in a Foreign Language, 15,* 1–27.

Watanabe, Y. (1997). Input, intake, and retention: Effects of increased processing on incidental learning of foreign language vocabulary. *Studies in Second Language Acquisition, 19,* 287–307.

Webb, S. (2007). The effects of repetition on vocabulary knowledge. *Applied Linguistics, 28,* 46–65.

Webb, S. (2009). The effects of pre-learning vocabulary on reading comprehension and writing. *The Canadian Modern Language Review, 65,* 441–470.

Webb, S., & Rodgers, M. P. H. (2009a). The vocabulary demands of television programs. *Language Learning, 59,* 335–366.

Webb, S., & Rodgers, M. P. H. (2009b). The lexical coverage of movies. *Applied Linguistics, 30,* 407–427.

Appendix A

Glossary for the Television Program Grey's Anatomy

ABDOMEN	EPI	MC	SEIZE
ANEURYSM	EPISODE	MCDREAMY	SEIZURE
AORTA	EVACUATE	MEDS	SHEPHERD
ARTERY	FATE	MER	SKULL
ASS	FETAL	MERCY	SOBER
BEAST	FEVER	MORON	SPINE
BELLY	FINN	MORPHINE	SQUAD
BIOPSY	FLUID	MRI	STALK
BOOB	FRACTURE	NEURO	SUCTION
BOWEL	FREAK	OVARY	SURGICAL
BULLET	FREAKING	PAGED	SUTURE
CELIBACY	GIRLFRIEND	PEE	SWEATER
CLAMP	GSW	PENIS	THANKSGIVING
CLOT	HORMONE	PILL	TRAILER
CONTRACTION	INTERN	PISS	TRANSPLANT
C-SECTION	JEALOUS	PORN	TRAUMA
CT	JUJU	PRE-OP	TUMOR
CUTE	KARMA	PROM	UTERUS

DEVICE	KISS	PSYCH	VALVE
DISCHARGE	KNIT	PULSE	VITALS
DISSECT	LIGHTNING	QUINTUPLETS	VOMIT
DOLL	LIQUOR	QUIT	WHISPER
DONOR	LIVER	RUPTURE	WHOA
DUDE	LVAD	SANTA	WHORE
ELEVATE	MAMMA	SATAN	X-RAY
ENEMA	MARSHAL	SCALPEL	YELL

Appendix B

Glossary for the Television Program House

ABNORMAL	DISORDER	LUPUS	RADIATE
ACUTE	DONOR	LYMPH	RASH
ALLERGY	EDEMA	MANIPULATE	RESPIRATOR
AMPHETAMINE	ELEVATE	MARROW	SANE
ANAPHYLAXIS	EPI	MEDICATION	SCAR
ANEMIA	EPO	MEDS	SCLEROSIS
ANTIBIOTIC	ERR	MENINGITIS	SEIZURE
ANTIBODY	ETHICAL	METAPHOR	SHRINK
ARTERY	FAKE	MIGRAINE	SINUS
ASS	FEVER	MORPHINE	SODIUM
AUTOIMMUNE	FLUID	MRI	SPINE
AUTOPSY	FUNGUS	MS	STAT
BACTERIUM	GENETIC	NAUSEA	STEROID
BENIGN	GIRLFRIEND	NEUROLOGICAL	SWELL
BIOPSY	HALLUCINATE	NEUROLOGIST	SYNDROME
BITCH	HALLWAY	NICU	TACHYCARDIA
BULLET	HEP	OBSESS	TOX
CALCIUM	HERPES	OXYGEN	TOXIC
CANE	ICU	PANCREAS	TOXIN
CARDIAC	IMMUNE	PANTS	TRANSPLANT
CLOSET	INDUCE	PARALYSIS	TRASH
CLOT	INFLAME	PARASITE	TRAUMA
COMA	INSULIN	PEE	TUBERCULOSIS
CONSENT	INTERFERON	PILL	TUMOR
CORTEX	JERK	PISS	TWITCH
CT	JERSEY	PNEUMONIA	ULCER
DIAGNOSE	KISS	PRESCRIBE	URINE
DIAGNOSIS	LESION	PROTEIN	VALVE
DIAGNOSTIC	LIVER	PSYCHOSIS	VIRAL
DIARRHEA	LOBE	PULMONARY	VOMIT
DIC	LP	PUNCTURE	WHOA
DIFFERENTIAL	LUMBAR	QUIT	X-RAY

Appendix C

Glossary for the Medical Genre

ABDOMEN	ELEVATE	MCDREAMY	SKULL
ABNORMAL	EPI	MEDICATION	SPINE
ALLERGY	ERR	MEDS	STEROID
ANEURYSM	FAKE	MORPHINE	SURGICAL
ANTIBIOTIC	FEVER	MRI	SWELL
AORTA	FLUID	MS	SYNDROME
ARTERY	FREAK	OVARY	THANKSGIVING
ASS	FREAKING	OXYGEN	TOX
BACTERIUM	GIRLFRIEND	PAGED	TOXIN
BIOPSY	HALLUCINATE	PANTS	TRAILER
BITCH	IMMUNE	PARALYSIS	TRANSPLANT
BOWEL	INDUCE	PEE	TRAUMA
BULLET	INTERN	PILL	TUBERCULOSIS
CARDIAC	JEALOUS	PISS	TUMOR
CLOT	JERK	PROM	ULCER
COMA	JUJU	PROTEIN	URINE
CT	KISS	PSYCH	VALVE
CUTE	KNIT	PULMONARY	VOMIT
DIAGNOSE	LIVER	PULSE	WHOA
DIAGNOSIS	LP	PUNCTURE	X-RAY
DISCHARGE	LUMBAR	QUIT	YELL
DISORDER	LUPUS	SANE	
DONOR	LVAD	SCAR	
DUDE	MARROW	SEIZURE	

About the Author

Stuart Webb is a senior lecturer and the MA program director in the School of Linguistics and Applied Language Studies at Victoria University of Wellington. His research interests include vocabulary studies, and extensive reading and listening. He has published in journals such as *Applied Linguistics, Language Learning, Studies in Second Language Acquisition,* and *TESOL Quarterly*. Correspondence concerning this article should be addressed to Stuart Webb, Victoria University of Wellington, School of Linguistics and Applied Language Studies, P.O. Box 600, Wellington 6012, New Zealand. E-mail: stuart.webb@vuw.ac.nz