

# Vocabulary Demands of Television Programs

Stuart Webb

Victoria University of Wellington

Michael P. H. Rodgers

Fukuoka University

This study investigated vocabulary coverage and the number of encounters of low-frequency vocabulary in television programs. Eighty-eight television programs consisting of 264,384 running words were categorized according to genre. Television shows were classified as either British or American and then put into the following genres: news, drama, situation comedy, older programs, children's programs, and science fiction. The results showed that knowledge of the most frequent 3,000 word families plus proper nouns and marginal words provided 95.45% coverage, and knowledge of the most frequent 7,000 word families plus proper nouns and marginal words provided 98.27% coverage. The vocabulary size necessary to gain 95% coverage of the different genres ranged from 2,000 to 4,000 word families plus proper nouns and marginal words; 5,000 to 9,000 word families plus proper nouns and marginal words to gain 98% coverage. The analysis also indicated that there was great variation in coverage between episodes. The results showed that there were relatively few encounters with low-frequency vocabulary. However, if learners knew the most frequent 3,000 word families and they watched at least an hour of television a day, there is the potential for significant incidental vocabulary learning.

**Keywords** television; coverage of text; listening comprehension; incidental vocabulary learning; corpus studies; vocabulary frequency

There is little doubt that television provides a massive amount of first language (L1) input. Research suggests that people watch television in their native language from 3 (Statistics Canada, 2004) to 7 (Literacy Trust, 2006) hours a day

---

The authors wish to thank Paul Nation, Victoria University of Wellington, for his helpful comments on an earlier draft of this article, and the anonymous reviewers, for their useful comments.

Correspondence concerning this article should be addressed to Stuart Webb, Victoria University of Wellington, School of Linguistics and Applied Language Studies, P.O. 600, Wellington 6012, New Zealand. Internet: [stuart.webb@vuw.ac.nz](mailto:stuart.webb@vuw.ac.nz)

and that they spend a much larger amount of time watching television than they do reading (Statistics Canada, 1998; U.S. Department of Labor, 2006). Surprisingly, there is relatively little research examining the relationship between vocabulary and television viewing in comparison to research focusing on the relationship between vocabulary and reading, nor do there appear to be any data on the television viewing habits of second language (L2) learners. One might suspect that in an English as a second language (ESL) setting, learners may often watch L2 television programs, whereas in an English as a foreign language (EFL) setting, learners would have little opportunity to watch L2 programs. Whether language learners watch television may depend on whether they have the vocabulary necessary to understand programs. Reading comprehension research indicates that although there are many factors such as background knowledge and the structure of sentences and paragraphs that may affect comprehension, vocabulary knowledge is the most influential factor (Laufer & Sim, 1985). If L2 learners have reached the receptive vocabulary size sufficient to comprehend the vocabulary in television programs, they may be motivated to watch L2 television. However, less proficient learners are not likely to watch programs that they cannot understand. The aim of this study is to determine the vocabulary demands of television programs. Factors such as coverage (the percentage of known words in a text) and the number of times that vocabulary at different frequencies are encountered may provide useful data for researchers, teachers, and learners on the extent to which television may be used for language learning. Specifically, this study may provide some indication of the target vocabulary size necessary for adequate comprehension of television programs. Determining the point at which learners may understand the vocabulary in television may provide both teachers and learners with a vocabulary learning goal, which, when reached, would allow television to become a valuable source of L2 input.

### **How Much Vocabulary Is Needed to Understand Television Programs?**

Coverage may indicate how much vocabulary is necessary for adequate comprehension to occur and whether learners may be able to guess words from context. Coverage is a valuable measurement because it provides a target vocabulary size, which, when reached, may indicate whether learners will be able to use or understand that discourse. Primarily research focused on coverage has looked at the coverage needed for reading comprehension.

There are no L1 or L2 studies examining the relationship between coverage and television comprehension. However, research investigating the coverage

necessary for reading and listening comprehension may provide some indication of how much vocabulary is necessary for adequate comprehension of television programs. L2 studies have differed in the amount of text coverage that is needed for adequate comprehension to occur. Estimates are 95% for reasonable comprehension of a text (Laufer, 1989), 98% for learners to read for pleasure (Hirsh & Nation, 1992), 98% for adequate unassisted reading comprehension (Hu & Nation, 2000), and 98% for ideal coverage of written text (Nation, 2006). Hu and Nation's (2000) research is the most comprehensive study of text coverage. They examined L2 learners' comprehension of a relatively easy fiction text with differing amounts of text coverage. The results indicated that at 90% text coverage, a small number of learners gain adequate comprehension; at 95% text coverage, more learners gain adequate comprehension but they are still a minority; and at 100% text coverage, most learners are able to understand the text. A regression analysis indicated that 98% coverage was needed for adequate comprehension without use of a dictionary or glossary. Hu and Nation also suggested that because the text in their study was relatively easy to understand, learners may need greater coverage of texts from different genres such as newspapers and academic texts. Hu and Nation's study is supported by L1 research that also found text coverage of 98–99% to be appropriate, with the difference in coverage dependent on the difficulty of the text (Carver, 1994).

Nation (2006) examined the vocabulary size needed for reading several different types of discourse using word lists developed from the British National Corpus. Table 1 shows the vocabulary size needed to reach 95% and 98% coverage (including proper nouns) for reading a novel, newspapers, a graded reader, as well as listening to conversation and watching a children's movie. His analysis showed a marked difference among the different discourse types. As should be expected, the smallest vocabulary size was needed to read the graded reader because it was the only discourse type that was specifically designed for language learners. *The Picture of Dorian Gray* is a Level 3 graded reader from the Oxford Bookworm series. With knowledge of the first 2,000 word families and the proper nouns in the text, coverage is close to 97%, and with knowledge of the first 3,000 word families and proper nouns, coverage is almost 99%. In contrast, the most frequent 3,000 word families accounted for 96.03% of the words in conversation in the Wellington Spoken Corpus, whereas the most frequent 7,000 word families accounted for 97.95% coverage. Nation's results were supported by a study of the CANCODE corpus, which found that the most frequent 3,000 word families accounted for a similar coverage of spoken discourse (95.91%) (Adolphs & Schmitt, 2003). Nation also found that the most

**Table 1** Percentage text coverage including proper nouns of different discourse types

Word list	Discourse type				
	A graded reader	A movie	Conversation	Newspapers	A novel
	<i>The Picture of Dorian Gray</i>	<i>Shrek</i>	Wellington Spoken Corpus	Section A LOB Corpus	<i>Lady Chatterley's Lover</i>
2,000	96.75	87.91	90.38	89.62	90.14
3,000	98.86		96.03		
4,000		96.74		95.39	95.06
5,000					
6,000					
7,000		98.08	97.95		
8,000				98.31	
9,000					98.22
Proper nouns	5.55	1.47	1.03	5.29	2.05

Adapted from Nation, I. S. P. (2006). "How Large a Vocabulary Is Needed for Reading and Listening?" By I. S. P. Nation, 2006, *Canadian Modern Language Review*, 63, pp. 59–82. Reprinted with permission.

frequent 4,000 word families plus proper nouns and marginal words accounted for 96.74% coverage of the children's movie *Shrek*. However, at the 2,000-word level, *Shrek* accounted for the smallest text coverage (88%) of all of the discourse types. Including proper nouns and marginal words, the most frequent 7,000 word families accounted for 98.08% of the running words in *Shrek*. The coverage was fairly similar for newspapers and novels, which contained a larger number of less frequent words than the other genres. Including proper nouns, the most frequent 4,000 word families accounted for 95% of the running words in both genres, and the most frequent 8,000 word families accounted for 98% of the running words in newspapers, while the most frequent 9,000 words families accounted for 98.22 of the running words in the novel *Lady Chatterley's Lover*. Nation concluded that at the ideal text coverage of 98%, a vocabulary of 8,000–9,000 word families is needed for reading, and a vocabulary of 6,000–7,000 word families is needed for spoken text.

In the only study examining the effects of coverage on listening comprehension, Bonk (2000) examined listening comprehension of four short passages of increasing lexical difficulty. He measured comprehension with a written recall

test and a dictation test. He concluded that with the effective use of coping strategies, learners can have adequate listening comprehension at far below 95% for short texts. In contrast, Nation (2006) suggested that coverage greater than 98% may be needed for spoken language, and Donkaewbua (2007) reported that learners may need 99% coverage for adequate comprehension when listening to graded readers.

Comprehension of television programs may be easier than listening comprehension of conversation or radio programs because the vocabulary heard in television programs may be supported by visual imagery. Mueller (1980) found that for less proficient language learners, a visual cue (a drawing) facilitated listening comprehension but had no effect with more proficient learners. Hanley, Herron, and Cole (1995) compared the effects of learning with video and learning with pictures. They found that learners who watched short video clips in the target L2 (French) scored higher on a comprehension test than learners who heard the instructor read a transcript of the video and saw four pictures relating to the context. They suggest that video improved comprehension more than pictures because it is more effective at creating the link between L2 form and L1 meaning. Rubin (1994) suggested that comprehension may vary within and between television programs. Some types of television programs may provide ample visual clues that may help viewer's comprehension, whereas others may not. Similarly, the imagery in programs at different points within a program may aid comprehension at some points, whereas at other points, it may not.

Taken together, the research indicates that for adequate comprehension of television programs, coverage is likely to range from 90% to 99%. Because 95% and 98% may represent reasonable (Laufer, 1989) and ideal (Nation, 2006) comprehension of written text, they were chosen as the possible lower and upper boundaries, which may indicate adequate comprehension of television programs for this study. It is also important to note that 95% and 98% coverage do not equal 95% and 98% comprehension. Coverage refers to the percentage of known words in the discourse. Knowing the words does not necessarily indicate that the discourse will be understood. Coverage may be the most important factor in determining comprehension, but it is one of many factors that are involved in comprehension. One aim of the present study was to determine the vocabulary needed to reach 95% and 98% coverage of television programs. This may provide some indication of the stage of vocabulary development at which learners could understand television programs and at which teachers may effectively use television programs in the classroom.

## Do People Incidentally Learn Vocabulary When Watching Television?

Several studies have investigated the relationship between watching television or video and incidental vocabulary learning. Incidental vocabulary learning may be defined as, “learning words without deliberate decision to commit information to memory” (Laufer & Hulstijn, 2001, p. 11). Rice and Woodsmall (1988) found that children who watched a 12-min cartoon with a narration that featured 20 unknown L1 target words had higher scores on a picture recognition task that measured knowledge of those target words than children who watched the cartoon with a narration in which the target words did not occur. The findings also showed that 5-year-olds learned more words than 3-year-olds. Rice and Woodsmall suggested that accumulated linguistic knowledge and prior viewing experience may have a positive influence on incidental vocabulary learning from television. They also attributed learning to the number of encounters with the target words. All of the target items were encountered at least five times in the video. Oetting, Rice, and Swank (1995) used the same video and target words, with normally developed and with specific-language-impaired 6- to 8-year-olds. They found that both groups of learners demonstrated significant L1 incidental vocabulary gains with the size of the gains related to age and learner aptitude. The specific-language-impaired learners made the smallest gains, whereas the oldest learners with normal language development made the largest gains. The results provide some indication that the learners who knew the most words and had the greatest coverage learned the most words.

Pavakanun and d’Ydewalle (1992) examined the effects of watching a 12-min cartoon on vocabulary learning. They found that two groups of language learners (English native speakers learning Dutch and Dutch native speakers learning Spanish), who watched the movie in the target language, had higher scores on a multiple-choice test measuring recognition of meaning. All of the target words were spoken in the program at least five times. They concluded that watching television programs in a foreign language may lead to large incidental vocabulary learning. Neuman and Koskinen (1992) compared L2 incidental vocabulary learning through reading, watching television, and watching television with subtitles with seventh- and eighth-grade students. They found that students who watched captioned television programs significantly outperformed those who read the text of the programs on three different measures of vocabulary knowledge. The captioned television group also had higher mean scores than the traditional television group but not all of the differences were significant. The traditional television group had higher scores than the reading group on

all of the tests, but the differences between the two groups were not reported. The results also indicated that the most proficient L2 students learned more vocabulary than the less proficient students.

In a study with foreign language learners, d'Ydewalle and Van de Poel (1999) found that 8- to 12-year-old children who watched a 10-min movie with successive static pictures made significant incidental L2 vocabulary gains. Target words were encountered in the movie at least four times. The results showed that learning was superior when the target language was presented in the soundtrack rather than in the subtitles. Koolstra and Beentjes (1999) examined the effects of watching television on foreign language vocabulary acquisition with children in primary school. They found that young children who watched a 15-min English language documentary with Dutch subtitles had higher scores on a multiple-choice test that involved matching the L1 meanings with the L2 target vocabulary than children who watched the program without subtitles and a control group who did not watch the program. Scores were also significantly higher in the nonsubtitled condition than in the control group, indicating that watching television did lead to incidental vocabulary learning. The results also showed that the children who frequently watched L2 television programs with L1 subtitles scored higher on the multiple-choice test than children who watched less often, indicating that increased viewing may lead to increased vocabulary learning. Koolstra and Beentjes reported that approximately half of the television programs that Dutch children watch are L2 programs with L1 subtitles and suggested that watching these shows may lead to large cumulative gains in vocabulary knowledge. They also suggested that programs that use "clear and grammatically correct language, visual aids for comprehension of word meanings, and subjects that kindle the children's interest may be stronger than the effects observed" in their experiment (p. 59).

One important question that was never addressed in any of the above studies on incidental learning is: What amount of text coverage is needed for incidental vocabulary learning to occur? Liu and Nation (1985) suggested that 95% coverage is necessary for learners to guess words from context in written text, and Nation (2001) suggested that 98% coverage is ideal for guessing words from context when reading. However, there is reason to believe that the combination of visual imagery and aural input may make it easier to learn words incidentally through watching television programs than through reading. Elley (1989) found that when children listened to stories, the chances of vocabulary learning were improved with illustrations. In a study involving students reading and listening to a graded reader, Horst, Cobb, and Meara (1998) suggested that the illustrations presented in the book were likely to aid vocabulary learning.

These studies suggest that visual input may increase the chances of incidental learning. Whether the necessary coverage is smaller for television programs in comparison to written text remains unknown and is not within the scope of this study. However, it is important for researchers to consider the effects of coverage on incidental learning because it could provide a gauge to whether learners of a particular vocabulary size may be able to learn vocabulary incidentally through watching television. Several studies found that older learners learned more words than younger learners (Neuman & Koskinen, 1992; Oetting et al., 1995; Rice & Woodsmall, 1988). This could be due to older learners having a larger vocabulary and higher coverage of the programs. Unless vocabulary coverage is measured, it is difficult to get an accurate assessment of learning.

Taken as a whole, the research indicates that watching television may lead to incidental vocabulary learning. However, it should be noted that in most cases, the television programs used were designed for children and provided explicit and redundant information, which increased the chances of vocabulary learning. Most studies also noted that there were several encounters with target words that would also increase the potential for learning (Nagy, Herman, & Anderson, 1985). Research has repeatedly shown that the chance of incidentally learning words tends to increase as the number of encounters with those words increases (Horst et al., 1998; Jenkins, Stein, & Wysocki, 1984; Rott, 1999; Saragi, Nation, & Meister, 1978; Waring & Takaki, 2003; Webb, 2007). Estimates have varied ranging from 6 (Rott) to 20 encounters (Waring & Takaki) needed to learn a word in written context. Thus, the learning conditions in the studies were highly conducive to vocabulary learning. This may not be the case with regular television programs. The vocabulary demands of television programs are likely to vary between genres and programs within a genre. At present, there is not any research that has examined the vocabulary demands of television programs. Research investigating coverage and the number of encounters with low-frequency words may provide some indication of the thresholds needed to reach adequate comprehension and to incidentally learn vocabulary. We therefore formulate the following research questions:

1. How many words do viewers need to know to understand the vocabulary in television programs?
2. How does coverage of British television differ between programs from different genres?
3. How does coverage of American television differ between programs from different genres?

4. How often do viewers encounter low-frequency words in television programs?

## Methodology

### Materials

Eighty-eight television programs were used in this study. Fifty-five of the programs were more traditional half-hour or 1-hr programs, and 33 were news stories. The total viewing time for all of the programs was approximately 35.1 hr. However, the amount of time viewers might spend in front of a television could be much longer because commercial breaks were not included in the total; for example, a drama such as *ER* runs over 60 min, but in this analysis, it was only listed as 44 min because there were 16 min of commercials. The length of time was unknown for a number of news programs. The time of those programs was calculated by dividing the number of running words by the number of words per minute in known news programs. The programs were selected according to the following factors: availability of transcripts, genre, and popularity of the programs. The aim of the selection process was to choose programs that fit a number of common television genres, were well known, and, therefore, were likely to have been watched by a large number of people. The programs were first classified as either originating from the United Kingdom or the United States and then classified according to genre. American programs were grouped into six genres: drama, situation comedy, science fiction, children's programs, older programs, and news. Because there were not many British television transcripts available, British television programs were grouped into two categories: British news and British programs. The British programs category included programs from all of the genres examined in the American television group. British drama, situation comedies, older programs, children's programs, and science fiction programs were classified as British programs. The title of each television program and episode is listed under its genre in the Appendix.

All of the words that could not be heard in the television programs, such as stage commands, storyline, and speakers' names, were removed from the transcripts. Because removing the inaudible language is a long and time-consuming process, the number of hours of television program was limited to approximately 35.1 hr. This represents approximately 11.3 days of viewing time (excluding commercials) for L1 viewers<sup>1</sup> and would likely represent a significantly larger number of days viewing for L2 viewers.

Contractions, connected speech, and hyphenated words were changed to conform with the spellings used in the word lists. This accounted for 0.47%

of the tokens in the study. For example, *Y'know*, *wanna*, *livin'*, *listenin'*, and *wouda* were changed to *you know*, *want to*, *living*, *listening*, and *would have*, respectively. If the spellings of the words were not changed, they would have been classified as words that were less frequent than the 14,000-word level.<sup>2</sup> Proper nouns that were not listed in List 15 (proper noun list) and were originally listed in *Not in the lists* were reclassified as proper nouns and added to List 15.

### Analysis

The transcripts were analyzed with the RANGE software (Nation & Heatley, 2002). This is a computer program that can list the word families that occur in a text according to their frequency. Fourteen 1,000-word-frequency lists were used with the RANGE software to show the 1,000-word level at which the words in the text occurred. Nation (2006) created the fourteen 1,000-word lists from their range and frequency of occurrence in the British National Corpus (BNC). The word families in the lists were rated as Level 6 according to Bauer and Nation's (1993) classification of word families. Level-6 word families include inflections and more than 80 derivational affixes. All word stems were free forms not bound forms. Less frequent words that were not found in the fourteen 1,000-word lists may have been classified by the RANGE program as proper nouns (List 15), marginal words such as *oh*, *uh*, *mmm*, and *ah* (List 16), or *Not in the lists*. It should be noted that several words such as *Internet* and *meow* that appeared in *Not in the lists* were likely to be known by learners with a vocabulary size of 2,000 word families. This would suggest that the coverage figures may be slightly conservative. The RANGE program and the words lists can be downloaded from Paul Nation's Web site: <http://www.victoria.ac.nz/lals/staff/paul-nation/nation.aspx>. For more information about the word lists, see Nation (2004, 2006).

### Results and Discussion

The television programs used 264,384 words and accounted for 5,888 of the 14,000 Level-6 word families from the BNC lists. Table 2 shows that the first 1,000 word families accounted for most of the tokens, types, and word families, with the number of each consistently decreasing as the word frequency decreases. The most frequent 1,000 word families accounted for 225,011 of the running words (85.11%). In contrast, the second set of 1,000 word families accounted for only 11,675 running words (4.42%). By the fifth set of 1,000 word families, the percentage of tokens from the text is only 0.71%, indicating the relative significance of knowing the most frequent words. However, it should

**Table 2** Tokens, types, and word families at each word level for all television programs

Word list	Tokens		Types		Word families
	Raw	%	Raw	%	
1,000	225,011	85.11	3,169	22.79	996
2,000	11,675	4.42	2,269	16.32	963
3,000	5,095	1.93	1,457	10.48	820
4,000	3,439	1.30	1,019	7.33	694
5,000	1,872	0.71	739	5.31	539
6,000	1,307	0.49	519	3.73	422
7,000	857	0.32	386	2.78	333
8,000	560	0.21	292	2.10	250
9,000	525	0.20	267	1.92	231
10,000	401	0.15	181	1.30	168
11,000	339	0.13	207	1.49	187
12,000	230	0.09	122	0.88	109
13,000	179	0.07	116	0.83	106
14,000	103	0.04	73	0.52	70
Proper nouns	7,827	2.96	1,656	11.91	1,656
Marginal words	2,719	1.03	43	0.31	4
Not in the lists	2,245	0.85	1,392	10.01	? <sup>a</sup>
Total	264,384		13,907		7,513

<sup>a</sup>The RANGE program is not able to calculate word families for words not in the lists.

also be noted that there were still a large number of word families (539) that occurred at the fifth 1,000-word level. The difference is that the number of times word families at the 5,000-word level were encountered is much less in comparison to the most frequent families. In the third column of Table 2, row 15 shows that the percentage of proper nouns was 2.96%, which accounted for the third highest percentage of tokens after the first and second 1,000 words. This clearly demonstrates the importance of being able to recognize proper nouns when watching television. In the third column of Table 2, row 16 shows the percentage of marginal words (e.g., ah, oh, huh) was 1.03%. If we assume that proper nouns and marginal words have a minimal learning burden, as Nation (2006) did in his analysis of discourse types, then we can add those figures to the text coverage needed to know each 1,000-word list. Table 3 shows the cumulative coverage for the tokens in the television programs with and without proper nouns and marginal words.

In answer to the first research question, without knowledge of the proper nouns and marginal words (PNAMW), the vocabulary necessary to reach 95%

**Table 3** Cumulative coverage, with and without proper nouns and marginal words, for all television programs

Word list	Coverage without proper nouns and marginal words	Coverage including proper nouns and marginal words
1,000	85.11	89.10
2,000	89.53	93.52
3,000	91.46	95.45 <sup>a</sup>
4,000	92.76	96.75
5,000	93.47	97.46
6,000	93.96	97.95
7,000	94.28	98.27 <sup>b</sup>
8,000	94.49	98.48
9,000	94.69	98.68
10,000	94.84	98.83
11,000	94.97	98.96
12,000	95.06 <sup>a</sup>	99.05
13,000	95.13	99.12
14,000	95.17	99.16
Proper nouns	2.96	
Marginal words	1.03	

<sup>a</sup>Reaching 95% coverage.

<sup>b</sup>Reaching 98% coverage.

coverage is 12,000 word families, and with a vocabulary of 14,000 word families, coverage is only 95.17%. In contrast, a vocabulary of 3,000 word families plus PNAMW provides 95.45% coverage, and with a vocabulary of 7,000 word families plus PNAMW, 98.27% of the words would be known. This means that if learners knew the most frequent 3,000 word families and PNAMW, about 1 in every 22 words would be unknown. If learners knew the most frequent 7,000 word families and PNAMW, about 1 out of every 58 words would be unknown. This would be around three to four unknown words per minute of continuous speech.

Table 4 shows the cumulative coverage for the tokens in the British and American television programs, with and without PNAMW. There were 193,460 running words in the American television programs and 70,924 running words in the British television programs. Including proper nouns and marginal words, the coverage at the 1,000-word level was 89.51% for British television and 88.95% for American television. Table 4 shows that a vocabulary of 3,000 word families plus PNAMW would provide 95.31% coverage and a vocabulary

**Table 4** Cumulative coverage, with and without proper nouns and marginal words, for British and American television

Word list	British		American	
	Coverage without proper nouns and marginal words	Coverage including proper nouns and marginal words	Coverage without proper nouns and marginal words	Coverage including proper nouns and marginal words
1,000	85.66	89.51	84.91	88.95
2,000	90.08	93.93	89.32	93.36
3,000	91.94	95.79 <sup>a</sup>	91.27	95.31 <sup>a</sup>
4,000	93.20	97.05	92.59	96.63
5,000	93.90	97.75	93.30	97.34
6,000	94.34	98.19 <sup>b</sup>	93.81	97.85
7,000	94.67	98.52	94.13	98.17 <sup>b</sup>
8,000	94.92	98.77	94.33	98.37
9,000	95.09	98.94	94.54	98.58
10,000	95.20	99.05	94.71	98.75
11,000	95.34	99.19	94.83	98.87
12,000	95.44	99.29	94.91	98.95
13,000	95.47	99.32	94.99	99.03
14,000	95.49	99.34	95.04	99.08
Proper nouns	3.09		2.91	
Marginal words	0.76		1.13	
Tokens	70,924		193,460	

<sup>a</sup>Reaching 95% coverage.

<sup>b</sup>Reaching 98% coverage.

of 7,000 word families plus PNAMW would provide 98.17% coverage of American television. In contrast, assuming that learners knew the PNAMW, a vocabulary of 3,000 word families would provide 95.79% coverage of British television and a vocabulary of 6,000 words would reach 98.19% coverage. This result suggests that British television is slightly less demanding than American television because it makes greater use of higher frequency vocabulary. However, it should be noted that because the BNC consists primarily of British text (Nation, 2004), it may be more likely to represent the language in British television than in American television. This can be seen in the following head-words taken from the 1,000- and 2,000-word-level lists: *bin*, *biscuit*, *bloke*,

**Table 5** Cumulative coverage, including proper nouns and marginal words for British television within a genre and episode

Word list	All British television	British news		British programs	
		All	One story	All	One program
1,000	89.51	86.27	80.59	90.47	88.37
2,000	93.93	92.26	89.53	94.43	93.48
3,000	95.79 <sup>a</sup>	94.61	91.83	96.15 <sup>a</sup>	95.40 <sup>a</sup>
4,000	97.05	96.19 <sup>a</sup>	94.41	97.31	96.81
5,000	97.75	97.21	96.37 <sup>a</sup>	97.91	97.40
6,000	98.19 <sup>b</sup>	97.83	96.65	98.30 <sup>b</sup>	97.62
7,000	98.52	98.51 <sup>b</sup>	97.28	98.52	97.90
8,000	98.77	98.90	97.84	98.72	98.27 <sup>b</sup>
9,000	98.94	99.06	97.84	98.89	98.46
10,000	99.05	99.20	97.91	98.99	98.61
11,000	99.19	99.35	98.05 <sup>b</sup>	99.12	98.76
12,000	99.29	99.46	98.12	99.21	98.93
13,000	99.32	99.48	98.19	99.24	98.99
14,000	99.34	99.54	98.19	99.25	99.02
Proper nouns	3.09	3.79	4.96	2.88	3.63
Marginal words	0.76	0.03	0.00	0.98	0.28
Tokens	70,924	16,213	1,432	54,711	7,245

<sup>a</sup>Reaching 95% coverage.

<sup>b</sup>Reaching 98% coverage.

*bugger, chap, crisp, fortnight, lorry, motorway, muck, nick, nil, nought, parish, parliament, pence, petrol, pint, pound, pub, pudding, quid, royal, rubbish, sack, and sod.* All of those words are likely to occur at less frequent levels in a predominantly American corpus. This suggests that there may be little difference between the vocabulary size needed to watch American and British programs, with coverage perhaps being dependent on the amount of discourse (American or British) in the corpus from which the frequency lists are derived. Further research comparing British and American discourse with frequency lists derived from American discourse would be a useful follow-up to this study.

Table 5 shows the cumulative coverage including PNAMW for British television within a genre and within a single program from each genre. The bottom row of Table 5 lists the number of tokens indicating the relative size of each genre and program. Combined, British television and news was 70,924 running words long and was made up of 16,213 words from British news and

54,711 words from British programs. The difference in size between news and British programs could be attributed to the relatively small number of tokens in each news story. The BBC news story *Has someone been sitting on the FBI?* was only 1,432 words in comparison to 7,245 words from the episode of the British program *Spooks* entitled *Thou shall not kill*. Typically, the news programs selected were only a few minutes long, whereas British programs were from 30 min to 1 hr long.

In answer to the second research question, the results indicated that coverage of British television varied slightly between genres but that there were large differences between episodes. The results also provided some evidence that British television programs such as British dramas, situation comedies, and science fiction programs may be less demanding than British news. A vocabulary of 3,000 word families and PNAMW reached 96.15% coverage of British programs, whereas knowledge of the most frequent 4,000 word families and PNAMW was needed to reach that level for British news. To reach 98% coverage, only the most frequent 6,000 word families and PNAMW were needed for British programs, whereas a vocabulary of 7,000 words plus PNAMW was needed for British news. The analysis of single episodes suggested that coverage of news stories may vary significantly. With a vocabulary of 5,000 word families and PNAMW, 96.37% of the words would be known in the single news story. With a vocabulary of 11,000 words and proper nouns, 98.05% of the words would be known. The large difference in coverage between the single news story and the British news genre (1,000 word families to reach 95% and 4,000 word families to reach 98%) may be due, in part, to the short length of news stories. It might also be due to the wide range of news topics with the vocabulary demands varying a great deal from story to story. The single episode of the British program examined in this study provided nearly identical results to the genre. Although the vocabulary demands of single programs are also likely to vary somewhat between episodes, the much greater length of an episode in comparison to a news story may help to provide coverage that is closer to the overall findings within a genre. Knowledge of the most frequent 3,000 word families provided 95% coverage for both the single episode and British programs, whereas knowledge of 6,000 and 8,000 word families were needed to reach 98% coverage for the single episode and the genre, respectively.

The cumulative coverage including proper nouns for American television within a genre and within a single program from each genre is shown in Table 6. The bottom row of Table 6 lists the number of tokens indicating the relative size of each genre and single programs. American television was 193,460 running words long. It was made up of 18,490 words from news programs, 53,922 words

**Table 6** Cumulative coverage including proper nouns and marginal words for American television within a genre and episode

Word list	News		Drama		Science Fiction		Children's		Older programs		Situation Comedies		
	American television	All	One story	All	One program								
	1,000	88.95	84.58	85.27	89.63	91.04	86.96	87.78	91.67	90.05	90.59	93.26	91.12
2,000	93.36	90.97	93.30	93.62	95.03 <sup>a</sup>	92.10	92.36	95.13 <sup>a</sup>	93.76	94.42	96.53 <sup>a</sup>	94.78	96.82 <sup>a</sup>
3,000	95.31 <sup>a</sup>	93.73	95.79 <sup>a</sup>	95.50 <sup>a</sup>	96.07	94.25	94.21	96.64	94.87	96.26 <sup>a</sup>	97.83	96.32 <sup>a</sup>	97.55
4,000	96.63	95.38 <sup>a</sup>	96.75	96.99	97.42	95.70 <sup>a</sup>	95.33 <sup>a</sup>	97.59	95.09 <sup>a</sup>	97.28	98.45 <sup>b</sup>	97.33	98.10 <sup>b</sup>
5,000	97.34	96.23	97.13	97.66	97.94	96.66	96.04	98.29 <sup>b</sup>	96.85	97.68	98.86	97.90	98.45
6,000	97.85	97.18	97.51	98.05 <sup>b</sup>	98.50 <sup>b</sup>	97.27	96.54	98.58	97.07	98.16 <sup>b</sup>	99.21	98.34 <sup>b</sup>	98.62
7,000	98.17 <sup>b</sup>	97.92	98.47 <sup>b</sup>	98.32	98.83	97.63	97.00	98.78	97.10	98.47	99.29	98.53	98.79
8,000	98.37	98.21 <sup>b</sup>	98.66	98.51	98.95	97.88	97.21	98.90	97.15	98.67	99.48	98.65	98.79
9,000	98.58	98.52	98.66	98.70	99.05	98.05 <sup>b</sup>	97.46	99.03	97.18	99.09	99.56	98.77	98.89
10,000	98.75	98.65	98.85	98.83	99.14	98.30	97.65	99.24	97.59	99.17	99.59	98.94	98.89
11,000	98.87	98.81	98.85	98.94	99.23	98.49	97.75	99.32	97.75	99.25	99.67	99.00	99.03
12,000	98.95	98.90	98.85	99.01	99.27	98.62	97.88	99.38	97.75	99.29	99.72	99.07	99.03
13,000	99.03	99.06	99.61	99.08	99.36	98.70	98.05 <sup>b</sup>	99.43	97.75	99.33	99.75	99.17	99.13
14,000	99.08	99.14	99.61	99.13	99.42	98.76	98.81	99.45	97.75	99.35	99.78	99.21	99.30
Proper nouns	2.91	3.52	2.29	3.23	3.87	2.30	1.91	3.37	2.00	2.98	2.79	2.75	5.48
Marginal words	1.13	0.09	0.00	0.68	0.25	0.54	0.76	1.81	2.38	1.98	0.97	2.49	3.64
Tokens	193,460	18,490	523	53,922	6,741	51,344	5,245	13,090	3,691	25,413	3,696	31,201	2,885

<sup>a</sup>reaching 95% coverage, <sup>b</sup>reaching 98% coverage

from dramas, 51,344 words from science fiction programs, 13,090 words from children's programs, 25,413 words from older programs, and 31,201 words from situation comedies. The difference in tokens among the genres may be attributed to episode length. Typically, dramas and science fiction programs were about 44 min, whereas children's programs, situation comedies, and older programs were usually 22 min in length. This is apparent when looking at the number of running words in a random episode from each genre. The bottom row shows that a single episode of a drama accounted for 6,741 words. In contrast, a single episode of a situation comedy accounted for 2,885 words. News stories were the shortest in length and accounted for the fewest words per episode. There were only 523 words in the CNN news story *Buildings collapse in San Francisco*.

In response to the third research question, the results indicated that the vocabulary necessary to reach 95% coverage was fairly consistent between American genres although there were large differences between episodes. Together, knowledge of the most frequent 3,000 word families and PNAMW was needed to reach 95% coverage of American television. Knowledge of the most frequent 3,000 word families and PNAMW also provided 95% coverage of dramas, older programs, and situation comedies. Only 2,000 word families and PNAMW were necessary to reach 95% coverage of children's programs, and 4,000 word families and PNAMW were necessary to reach 95% coverage of news stories and science fiction programs. There were larger differences in the vocabulary necessary to reach 98% coverage between the American genres. A vocabulary of the most frequent 5,000 word families plus PNAMW was sufficient to reach 98% coverage of children's programs; a vocabulary of the most frequent 6,000 word families and PNAMW reached 98% coverage of dramas, older programs, and situation comedies; a vocabulary of the most frequent 8,000 and 9,000 word families and PNAMW reached 98% coverage of news stories and science fiction programs, respectively. The analysis indicated that at both 95% and 98% coverage, children's programs were the least demanding, needing the smallest vocabulary sizes to reach those coverage points.

The analysis also indicated that there may be great variation between episodes within a genre; for example, a vocabulary of 2,000 word families and PNAMW provided 95% coverage, and knowledge of 5,000 word families and PNAMW provided 98% coverage of children's programs. However, it was necessary to know 4,000 word families and PNAMW to reach 95% coverage and more than 14,000 word families and PNAMW to reach 98% in the episode randomly chosen from children's programs. This discrepancy was not only limited to children's programs. There were differences in the vocabulary necessary

to reach either 95% or 98% coverage between a single episode and its genre in all genres. This shows the importance of considering the vocabulary demands of television programs in order to determine whether a program is suitable for use in the classroom. It is particularly important that both teachers and learners are aware that the vocabulary demands vary between episodes. Learners may be discouraged from watching television if the episodes that they watch contain a large number of low-frequency words. However, if coverage is sufficient for comprehension, then learners might be more motivated to watch television outside of the classroom, which may lead to incidental vocabulary learning.

It is also important to note the large differences between reaching 95% and 98% coverage. For all of the programs in the study, the difference between 95% and 98% coverage was 4,000 word families. For many language students, learning the 3,000 most frequent word families needed to reach 95% coverage of television may be a long and difficult task. However, it is an achievable goal in both an EFL and ESL learning context. If 98% coverage is the threshold necessary for adequate comprehension and incidental learning, this may be an unattainable goal for many learners. This would be a particularly difficult vocabulary learning goal for EFL learners who get very little L2 input outside of the classroom. Only advanced L2 learners are likely to know the 7,000 word families that are needed to understand 98% of the vocabulary in most television programs.

In answer to the fourth research question, the results indicate that the number of encounters with most low-frequency word families is very small within 35.1 hr of viewing time; 55% of the low-frequency words were only encountered once and 19% were encountered twice, suggesting that with limited viewing over a variety of genres, incidental vocabulary learning is unlikely to occur for the majority of low-frequency words. There is, however, some evidence indicating the potential of incidental learning through television. Despite the relatively small number of programs in the study in comparison to the amount of time people watch television in their L1, 12% of the low-frequency words were encountered at least five times. This may be enough encounters with unknown words to at least gain partial vocabulary knowledge. If learners were to watch television regularly over a long period of time, the potential for learning would increase as the number of encounters with unknown and partially known words increased. Tables 7 and 8 show the number and percentage of encounters with low-frequency words (word families from the 5,000–14,000 word lists) that occurred within all of the television programs and within each genre.

Because there are fewer occurrences within each specific genre, the number and percentage of encounters are lower than in the entire television corpus. The high percentage of word families that were encountered only once (77%) and the

**Table 7** Number and percentage of encounters with 5,000–14,000-level word families

	All television		American news		British news		British programs	
	Amount	%	Amount	%	Amount	%	Amount	%
1 encounter	1,332	55%	225	69%	180	76%	465	66%
2 encounters	462	19%	47	14%	17	7%	115	16%
3–4 encounters	339	14%	32	10%	22	9%	65	9%
5–7 encounters	149	6%	11	3%	5	2%	30	4%
8–9 encounters	42	2%	3	1%	2	1%	4	1%
10+ encounters	91	4%	8	2%	12	5%	24	3%
Total word families	2,415		326		238		703	

low percentage of word families that were encountered five or more times (3%) within the situation comedies genre suggests that it may be the most difficult type of program in which to learn vocabulary incidentally. In contrast, the lower percentage of word families encountered once (61%) and the higher percentage of word families encountered five or more times (6%) indicate that incidental vocabulary learning is more likely to occur within children's programs. This should not be surprising, as one of the aims of children's programs used in this study is to promote vocabulary learning. It should also be noted that the corpus for the children's genre was quite small due to a lack of available transcripts. A larger children's corpus may show that the number of encounters with word families is much higher than other genres and most conducive to vocabulary learning. Table 7 also shows that the percentage of word families encountered at least five times was relatively high in both British (8%) and American (6%) news programs despite a smaller number of running words than most of the genres. This finding parallels those of Gardner (2004), who found that expository text tended to have a greater number of encounters with low-frequency words than narratives, which tended to use a greater proportion of high-frequency words. This suggests that television news could provide a greater opportunity for incidental learning than other programs. Moreover, if viewers were to follow news stories relating to particular events that are given a large amount of airplay, such as wars, elections, and sports leagues, it is likely that the number of encounters with low-frequency words within those storylines would be even higher and increase the chances of incidental learning.

The data for science fiction programs and older programs also suggests that they may be better suited to incidental learning than some other genres. Only

**Table 8** Number and percentage of encounters with 5,000–14,000-level word families

	Drama		Children		Comedy		Older programs		Science fiction	
	Amount	%	Amount	%	Amount	%	Amount	%	Amount	%
1 encounter	448	67%	58	61%	290	77%	152	68%	465	66%
2 encounters	126	19%	19	20%	53	14%	32	14%	115	16%
3–4 encounters	66	10%	12	13%	27	7%	25	11%	65	9%
5–7 encounters	16	2%	2	2%	4	1%	10	4%	30	4%
8–9 encounters	2	0.20%	1	1%	2	1%	0	0%	4	1%
10+ encounters	9	1%	3	3%	3	1%	6	3%	24	3%
Total word families	667		95		379		225		703	

**Table 9** Number and percentage of encounters with 4,000-level word families

	Television corpus	
	Amount	Percentage
1 encounter	224	32%
2 encounters	132	19%
3–4 encounters	135	19%
5–7 encounters	103	15%
8–9 encounters	25	4%
10+ encounters	75	11%
Total word families	694	

66% of the low-frequency word families that occurred were encountered once in science fiction programs and 68% were encountered once in older programs, whereas 8% of the word families occurred more than five times for science fiction programs and 7% for older programs.

Whereas the discussion above focused on the percentage of encounters with word families from the most frequent 5,000–14,000 words, perhaps the most likely words to be learned by viewers are at the next word level. In this case, if learners only knew the most frequent 3,000 word families, then the next word level would be the 4,000 level. Table 9 shows the number and percentage of encounters with the items from the 4,000-word list. At this level, there is much stronger evidence that television provides an opportunity for vocabulary learning. Only 32% of the word families at this level were encountered once and 30% were encountered at least five times. This suggests that for learners who know most of the most frequent 3,000 words, television may offer the potential for large incidental vocabulary gains.

The results showed that knowledge of the most frequent 3,000 word families plus PNAMW provided more than 95% coverage of television programs and knowledge of the most frequent 7,000 word families and PNAMW provided 98% coverage. We suggest that teachers and learners target the 3,000-word level as the point at which television might become a useful tool for language learning. There are several reasons for this. First, research indicates that the combination of visual and aural input might make comprehension of television programs easier than listening comprehension (Hanley et al., 1995; Mueller, 1980). Second, research suggests that watching television is at least as effective at promoting vocabulary knowledge as learning solely through written input (Neuman & Koskinen, 1992). Third, learning the most frequent 3,000 word families may be an attainable goal for many L2 learners. However, learning the

most frequent 7,000 words may not be a realistic goal for the majority of L2 learners. Whereas 98% coverage could be the ideal coverage, it is likely that coverage as low as 90% may also facilitate vocabulary learning and comprehension. However, both comprehension and learning are likely to improve with higher coverage.

Because people enjoy watching television and spend such a large amount of time doing so and because research indicates that watching television can lead to incidental vocabulary learning, television could provide a means of increasing vocabulary size. Research has shown that viewers who watch L2 television may have better comprehension of television (Secules, Herron, & Tomasello, 1992) and increased vocabulary learning from television (Koolstra & Beentjes, 1999) than viewers who watch fewer programs. Although we strongly advocate learning vocabulary through reading, the fact that people spend more time watching television than reading suggests that it could also be an effective method of learning vocabulary. This was particularly apparent at the 4,000-word level, at which 49% of word families were encountered at least three times. Research has shown that the more words are encountered in context, the more likely they will be learned (Horst et al., 1998; Jenkins et al., 1984; Nation, 2001; Rott, 1999; Waring & Takaki, 2003; Webb, 2007). Although three encounters may not lead to full knowledge of a word, it may lead to partial knowledge of several aspects of vocabulary knowledge (Webb). Watching for an average of 1 hr a day for 1 month would provide close to the amount of television analyzed in this study. Increased viewing would increase the likelihood of incidental learning. Watching an hour or more of television per day over 1 year could lead to large incidental learning gains. This should be possible in an ESL context, but it is probably quite rare in an EFL context, in which access to L2 television programs may be limited. It is important to note that the results suggest that there is likely to be very little incidental vocabulary learning from watching a single episode of a television program. This is because there may not be enough encounters to learn unknown words. If the aim is for learners to learn vocabulary from single episodes, then repeated viewings of a program are essential and may have a positive effect on vocabulary learning (Elley, 1989; Horst & Meara, 1999).

If 95% coverage is sufficient for comprehension, the results indicate that children's programs may be the least demanding. The most frequent 2,000 word families accounted for 95% coverage of children's programs. This should not be surprising, as children's programs aim to be easily understood. Using children's programs in the classroom initially may help to determine whether learners may be able to understand more demanding genres. This may also be a useful way to

focus learners on learning vocabulary through watching television. If learners are able to easily understand children's programs and know the most frequent 3,000 word families, American dramas, older programs, situation comedies, and British programs may be the most appropriate genres. The results indicated that both American and British news stories and science fiction programs may be the most demanding genres. This should not be too surprising because those programs are more likely to contain technical vocabulary than the other genres analyzed in this study. It is important to note that coverage is likely to vary from episode to episode and teachers and learners should not always expect coverage to be the same for each program. In an EFL context in which viewing may be limited, it is particularly important for teachers to consider the vocabulary demands of programs. The results showed that the vocabulary size necessary to gain 95% coverage of some episodes may be as low as the most frequent 2,000 word families, whereas it may be as high as 5,000 word families in others. Teachers may need to evaluate a number of programs before they are able to find one that is appropriate for their learners. Randomly selecting episodes is likely to lead to poor comprehension. When there is some uncertainty about the vocabulary demands of programs, preteaching low-frequency vocabulary and providing some background information may help to facilitate comprehension.

The results showed that although the majority of low-frequency word families were only encountered once (55%), there were still many low-frequency words that were encountered more than five times (12%). At the 4,000-word level, the number of word families encountered only once was much smaller (32%) and the number of word families encountered five or more times was much higher (30%). This suggests that if learners know the majority of the most frequent 3,000 word families, there is the potential for large incidental vocabulary learning at the next word level.

A wide range of television programs was examined in this study. However, it might be easier to understand programs from within a subgenre in which topics and storylines may be similar. Researchers and teachers should consider if there might be a genre- or subgenre-specific vocabulary; for example, police or crime-related shows may have a number of words that are specific to that subgenre. If learners were to specifically watch one subgenre, such as dramas relating to medicine or crime, the number of encounters with some low-frequency words may be much higher. One approach that teachers could take would be to preteach those words. Preteaching could be one method of improving comprehension and increasing vocabulary learning. Investigating whether there was a vocabulary specific to a genre or subgenre would be a useful follow-up to this study.

The results indicate that there is a need for further research investigating the relationship between watching television and vocabulary. In particular, several questions remain to be answered. First, what is the coverage necessary for adequate comprehension of television programs? Second, what is the coverage necessary for incidental vocabulary learning from watching television? Although we suggested 95% coverage based on research findings investigating text coverage and research examining learning with visual cues, there is still a need to examine the effects of different coverage points ranging from 90% to 99% experimentally through a controlled treatment to determine a minimum and ideal coverage. This would create a vocabulary learning goal for teachers and learners.

### Limitations

First, it should be noted that although the 35.1 hr of television analyzed may represent a significant number of days watching L2 television, it does not represent very many days watching L1 television. If learners were able to watch half as much L2 television as they do L1 television, they would have many more encounters with low-frequency words, resulting in increased learning. This is more likely to occur in an ESL setting with intermediate and advanced learners, who have a large enough vocabulary to reach 95% coverage.

Second, researchers, teachers, and learners need to be aware that vocabulary is one of many factors that contribute to comprehension. Factors such as background knowledge (Stahl, Hare, Sinatra, & Gregory, 1991; Stahl & Jacobson, 1986; Stahl, Jacobson, Davis, & Davis, 1989), the importance of the unknown vocabulary in the context (Stahl, 1990), the amount of redundant information (Kameenui, Carnine, & Freschi, 1982), and individual differences (Mezynski, 1983; Stahl) have been found to affect reading comprehension and may also affect comprehension of television programs. The correlation among meaning and imagery, the rate of speech, and the clarity of speech may also affect listening comprehension (Rubin, 1994). Although coverage may be a very important factor in comprehension of television programs, it is only one of a number of factors that need to be considered in studies of comprehension.

Revised version accepted 26 May 2008

### Notes

- 1 This is using Statistics Canada's (2004) data reporting 3.1 hr per day as the number of hours that people 2 years and older watch television.

- 2 Although the percentage of contractions and connected speech that were changed to conform with spellings from the BNC lists was very small (0.47%), it should be noted that knowing the changed spellings does not ensure that you would recognize those items in their original forms; for example, learners may know *would have* and *listening* but that does not necessarily mean they would recognize *woulda* and *listenin'*. It is likely that the amount of connected speech may be one of a number of factors (see Limitations section) that may affect comprehension and incidental vocabulary learning.

## References

- Adolphs, S., & Schmitt, N. (2003). Lexical coverage of spoken discourse. *Applied Linguistics*, 24, 425–438.
- Bauer, L., & Nation, I. S. P. (1993). Word families. *International Journal of Lexicography*, 6(4), 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.
- Donkaewbua, S. (2007). *The effects of previous partial word knowledge on vocabulary learning through listening*. Unpublished doctoral dissertation, Victoria University of Wellington, New Zealand.
- Duquette, L., & Painchaud, G. (1996). A Comparison of vocabulary acquisition in audio and video contexts. *Canadian Modern Language Review*, 54, 143–172.
- 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.
- Elley, W. B. (1989). Vocabulary acquisition from listening to stories. *Reading Research Quarterly*, 24, 174–187.
- Gardner, D. (2004). Vocabulary input through extensive reading: A comparison of words found in children's narrative and expository reading materials. *Applied Linguistics*, 25, 1–37.
- Hanley, J., Herron, C., & Cole, S. (1995). Using video as advance organizer to a written passage in the FLES classroom. *Modern Language Journal*, 79, 57–66.
- Hirsh, D., & Nation, P. (1992). What vocabulary size is needed to read unsimplified texts for pleasure? *Reading in a Foreign Language*, 8(2), 689–696.
- 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.
- Horst, M., & Meara, P. (1999). Test of a model for predicting second language lexical growth through reading. *Canadian Modern Language Review*, 56, 308–328.

- Hu, M., & Nation, I. S. P. (2000). Vocabulary density and reading comprehension. *Reading in a Foreign Language*, 13(1), 403–430.
- 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, UK: Multilingual Matters.
- Laufer, B., & Hulstijn, J. (2001). Incidental vocabulary acquisition in a second language: The construct of task-induced involvement. *Applied Linguistics*, 22, 1–26.
- Laufer, B., & Sim, D. D. (1985). An attempt to measure the threshold of competence for reading comprehension. *Foreign Language Annals*, 18, 405–411.
- Literacy Trust. (2006). One in 10 watches seven hours of TV a day but newspapers still popular. London: Author. Retrieved October 21, 2007, from <http://www.literacytrust.org.uk/database/tv.html>
- Liu, Na, & Nation, I. S. P. (1985). Factors affecting guessing vocabulary in context. *RELC Journal*, 16, 33–42.
- 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. *Modern Language Journal*, 64, 335–340.
- Nagy, W. E., Herman, P., & Anderson, R. C. (1985). Learning words from context. *Reading Research Quarterly*, 20, 233–253.
- Nation, I. S. P. (2001). *Learning vocabulary in another language*. Cambridge: Cambridge University Press.
- Nation, I. S. P. (2004). A study of the most frequent word families in the British National Corpus. In P. Bogaards & B. Laufer (Eds.), *Vocabulary in a second language: Selection, acquisition, and testing* (pp. 3–13). Amsterdam: Benjamins.
- Nation, I. S. P. (2006). How large a vocabulary is needed for reading and listening? *Canadian Modern Language Review*, 63, 59–82.
- Nation, I. S. P., & Heatley, A. (2002). Range: A program for the analysis of vocabulary in texts [software]. Available from <http://www.victoria.ac.nz/lals/staff/paul-nation/nation.aspx>
- 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(1), 95–106.

- Oetting, J. B., Rice, M. L., & Swank, L. K. (1995). Quick incidental learning (QUIL) of words by school-age children with and without SLI. *Journal of Speech and Hearing Research, 38*, 434–445.
- 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.
- Rice, M. L., & Woodsmall, L. (1988). Lessons from television: Children's word learning when viewing. *Child Development, 59*, 420–429.
- 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. *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. *Modern Language Journal, 76*, 480–490.
- Stahl, S. A. (1990). *Beyond the instrumentalist hypothesis: Some relationships between word meanings and comprehension* (Tech. Rep. 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.
- Statistics Canada. (2004). *Television viewing, by age and sex, by province*. Ottawa, Ontario: Author. Retrieved October 21, 2007, from <http://www40.statcan.ca/101/cst01/arts23.htm>
- Statistics Canada. (1998). *Average time spent on activities, by sex*. Ottawa, Ontario: Author. Retrieved October 21, 2007, from <http://www40.statcan.ca/101/cst01/famil36a.htm>
- U.S. Department of Labor. (2006). *American time use survey summary*. Washington, DC: Author. Retrieved October 21, 2007, from <http://www.bls.gov/news.release/atus.nr0.htm>
- Waring, R., & Takaki, M. (2003). At what rate do learners learn and retain new vocabulary from reading a graded reader? *Reading in a Foreign Language, 15*, 1–27. Retrieved September 5, 2006, from <http://nflrc.hawaii.edu/rfl/October2003/waring/waring.html>

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

## Appendix

### Television Programs Examined

American television: News

#### 1. ABC

<i>Episode Name</i>	<i>Episode Number</i>
Tornado Rips Through School	03/02/2007
Senator John McCain Running For President	03/01/2007
Fashion Statement	02/26/2007
A Marine's Progress	02/28/2007
Vets Fight For Health Care	02/28/2007

#### 2. CBS

<i>Episode Name</i>	<i>Episode Number</i>
TXU Agrees To Go Green After Private Equity Buyout	02/26/2007
Taliban Claims Dick Cheney Was Target Of Attack In Afghanistan	02/26/2007
Eric Nerhus On Surviving Shark Attack	02/22/2007
Back Scatter X-Ray Machine Debuts In Phoenix Today	02/23/2007
Honeybees Are Being Trained To Sniff Out Explosives, Other Dangerous Chemicals	02/26/2007

#### 3. CNN News

<i>Episode Name</i>	<i>Episode Number</i>
Taliban Comeback in Afghanistan	02/23/2007
Kidnapping Suspect Still on the Loose in Florida	02/24/2007
Buildings Collapse in San Francisco	02/27/2007
Market Rebounds from Big Losses	02/28/2007
More Severe Weather Expected in Much of Country	03/01/2007

#### 4. Fox News

<i>Episode Name</i>	<i>Episode Number</i>
Navy's Marine Mammal Program May Deploy Dolphins to Base	02/13/2007
Prince Harry Prepares for Deployment to Iraq	02/19/2007
British Pull Thousands of Troops Out of Iraq	02/21/2007
Democrats Attempt to Revise Iraq War Authorization	02/23/2007
Anna Nicole Saga Captivates America	02/24/2007

## 5. NBC News

<i>Episode Name</i>	<i>Episode Number</i>
Insurance Giant State Farm To Write No More Policies	02/14/2007
Parade Magazine's Rosalind Wiseman Speaks About Cyberbullying	02/22/2007
Hillary Clinton And Barack Obama Clearing Air About Comments Made By David Geffen	02/22/2007
Furor Over Squalid Living Conditions For Recovering Soldiers At Walter Reed Hospital	02/23/2007
Students At Baghdad University Attacked By Suicide Bomber	02/25/2007

## American television: Children's programs

## 6. "Fraggle Rock"

<i>Episode Name</i>	<i>Episode Number</i>
All Work and All Play	208
A Friend in Need	210

## 7. "Mr. Rogers"

<i>Episode Name</i>	<i>Episode Number</i>
When parents go to work 3	1613
When parents go to work 4	1614

## 8. "Sesame Street"

<i>Episode Name</i>	<i>Episode Number</i>
Death on the Street	11/24/1983

## American television: Dramas

## 9. "The West Wing"

<i>Episode Name</i>	<i>Episode Number</i>
The Short List	109
Swiss Diplomacy	409

## 10. "The Sopranos"

<i>Episode Name</i>	<i>Episode Number</i>
D-Girl	205
The Knight in White Satin Armor	212

## 11. "ER"

<i>Episode Name</i>	<i>Episode Number</i>
The Miracle Worker	510
Be Patient	617

## 12. "C.S.I."

<i>Episode Name</i>	<i>Episode Number</i>
Too Tough To Die	116
Getting Off	416

## 13. "24"

<i>Episode Name</i>	<i>Episode Number</i>
700 am to 800 am	108
700 pm to 800 pm	307

## American television: Older programs

## 14. "Gilligan's Island"

<i>Episode Name</i>	<i>Episode Number</i>
Two on a Raft	101
Home Sweet Hut	102

## 15. "I Love Lucy"

<i>Episode Name</i>	<i>Episode Number</i>
The Freezer	129
Lucy Does a TV Commercial	130

## 16. "The Fugitive"

<i>Episode Name</i>	<i>Episode Number</i>
Fear in a Desert City	101
Wife Killer	317

## 17. "The Monkees"

<i>Episode Name</i>	<i>Episode Number</i>
The Monstrous Monkee Mash	216
The Devil and Peter Tork	218

## American television: Science fiction

## 18. "The X-Files"

<i>Episode Name</i>	<i>Episode Number</i>
Pilot	101
Avatar	321

## 19. "Stargate SG-1"

<i>Episode Name</i>	<i>Episode Number</i>
Children of the Gods	101
Matter of Time	216

## 20. "Star Trek Voyager"

<i>Episode Name</i>	<i>Episode Number</i>
Caretaker	101
False Profits	305

21. "Firefly"	
<i>Episode Name</i>	<i>Episode Number</i>
Out of Gas	107
Trash	112
22. "Battlestar Galactica"	
<i>Episode Name</i>	<i>Episode Number</i>
33	101
Lay Down Your Burdens 2	220
American television: Situation comedies	
23. "Everybody Loves Raymond"	
<i>Episode Name</i>	<i>Episode Number</i>
I Love You	102
Turkey or Fish	110
24. "Frasier"	
<i>Episode Name</i>	<i>Episode Number</i>
The Good Son	101
The Perfect Guy	517
25. "Friends"	
<i>Episode Name</i>	<i>Episode Number</i>
The One With Joeys Fridge	619
The One Where the Monkey Gets Away	119
26. "Seinfeld"	
<i>Episode Name</i>	<i>Episode Number</i>
The Suzie	815
The Apology	909
27. "Will and Grace"	
<i>Episode Name</i>	<i>Episode Number</i>
Pilot	101
Board Games	712
British television: News	
28. BBC	
<i>Episode Name</i>	<i>Episode Number</i>
Relatives still waiting to hear why the Kursk sank	07/08/2001
Has someone been sitting on the FBI	06/11/2001
Saddam Hussein	10/04/2002
Rescued Dogs	05/24/2006
Real Spooks	30/04/2007

## 29. Sky News

<i>Episode Name</i>	<i>Episode Number</i>
Blair, Cameron & Campbell on climate change as part of its Green Britain	01/09/2007
Lunchtime Live with Kay Burley	01/26/2007
Interview with Blur drummer and Labour candidate Dave Rowntree	04/29/2007

## British television: British programs

## 30. "Red Dwarf"

<i>Episode Name</i>	<i>Episode Number</i>
The End	101
Only The Good	808

## 31. "Spooks"

<i>Episode Name</i>	<i>Episode Number</i>
Thou Shall Not Kill	101
Smoke and Mirrors	210

## 32. "Teachers"

<i>Episode Name</i>	<i>Episode Number</i>
Substitute	101
Unknown	302

## 33. "The Story of Tracy Beaker"

<i>Episode Name</i>	<i>Episode Number</i>
Moving On	V19
The Wedding	V20

## 34. "Lovejoy"

<i>Episode Name</i>	<i>Episode Number</i>
The Firefly Cage	101
The Judas Pair	105

## 35. "Eastenders"

<i>Episode Name</i>	<i>Episode Number</i>
June 22 1995	198
June 27 1995	200