

Article**Cartoon Word Lists for Second Language Learning Based on Extensive Viewing of Cartoons**

Kohei Kanayama*

Hokkaido University of Education, Japan

(Corresponding author. Email: kanayama.kohei@a.hokkyodai.ac.jp)

Akira Iwata

Hokkai-Gakuen University, Japan

Received: 15 April 2026; Accepted: 26 September 2026; Published: 19 October 2026

<https://doi.org/10.58304/tc.261016>**Abstract**

This study created *Cartoon Word Lists* (CWLs) for second language learning through extensive viewing of cartoons. We built the *English Cartoon Corpus* (ECC), consisting of 8,016,916 words from a total of 3,756 episodes streamed on Netflix. The ECC comprised four subcorpora grouped by recommended viewing age (i.e., toddlers, young children, older children, and adults). Then, the words that met the following three conditions were selected: (1) used 30 times or more per million words; (2) having a dispersion (Juilland's D) value of 0.6 or higher; and (3) not listed in the 2,809 flemmas from the *New General Service List* version 1.2 (NGSL 1.2, Browne et al., 2024). This resulted in the extraction of 271, 266, 137, and 235 flemmas in the CWLs for Toddlers, Young Children, Older Children, and Adults, respectively. We calculated the lexical coverage in the scripts of other cartoons that were not included in the ECC. The results showed that the NGSL 1.2 alone covered over 90% of the corpus, while the addition of the CWLs increased coverage to over 95%, indicating a substantial gain in lexical coverage for cartoon viewing.

Keywords

Word list, extensive cartoon viewing, audiovisual input, second language learning, lexical coverage

Introduction

One of the major problems for second language (L2) learners in an English as a foreign language (EFL) environment is that they have limited exposure to English outside the classroom, due to the lack of spoken input in their daily lives (Dang et al., 2022). One way to overcome this issue is using authentic viewing materials (e.g., TED Talks, YouTube) to make learning both engaging and enjoyable (Teng, 2021). Previous studies have shown that L2 learners can acquire vocabulary incidentally through audiovisual input (Kurokawa et al., 2025; Webb et al., 2023). According to the cognitive theory of multimedia learning (Mayer & Fiorella, 2022), learners process information more deeply when verbal and visual information are presented together than when verbal information is presented alone. Because cartoons present spoken language together with rich visual information, they may provide favorable conditions for incidental vocabulary learning. Moreover, several studies have shown that viewing videos can improve listening skills (Masrai, 2020; Rodgers & Webb, 2017), with more recent research indicating that viewing cartoons enhances listening, as long as L2 learners understand the content (Kanayama, 2024; 2026). Thus, the present study proposes the extensive viewing of

cartoons as a means of continuous English learning. However, little attention has been paid to developing cartoon-specific word lists to support comprehension during cartoon viewing. Indeed, it remains unclear whether existing word lists adequately cover the words used in cartoons. Because understanding cartoons depends on learners' receptive knowledge of the words they contain, a domain-specific word list may facilitate learners' comprehension during cartoon viewing and thereby support L2 learning through extensive viewing. Therefore, we created *Cartoon Word Lists* (CWLs).

Literature Review

Creating word lists

L2 learners should learn high-frequency words first, regardless of their learning purposes. In this case, the *General Service List* (GSL) (West, 1953) is the oldest, but the most influential high-frequency word list. However, since the corpus used for GSL was created in the 1930s, it contains many outdated words (e.g., oar, spade, telegraph, milkmaid). In other words, the GSL has long been in need of revision (Brezina & Gablasova, 2015; Browne, 2014). In recent years, two new *General Service Lists* have been created to adapt the original GSL to current English usage. Brezina and Gablasova (2015) created the new-GSL (2,494 lemmas) from four large corpora (i.e., the *Lancaster-Oslo-Bergen Corpus*, the *British National Corpus* [BNC], the *BE06 Corpus of British English* [BE06], and the *EnTenTen 12*), which contain approximately 12 billion running words. The words in the NGSL, developed by Browne (2013), were extracted from 273 million running words from nine corpora (i.e., learner, fiction, journal, magazine, non-fiction, radio, spoken, documents, and television) of the *Cambridge English Corpus* (CEC). The first NGSL version 1.0 (NGSL 1.0) contained 2,818 lemmas, while the latest version (NGSL 1.2) consists of 2,809 lemmas.

After acquiring high-frequency words, L2 learners should focus on words that match their own learning purposes. Regardless of the texts, learners should know at least 95% of the words, without the help of a dictionary or other resources (Hu & Nation, 2000). One of the most efficient ways to increase the lexical coverage in a particular domain is to learn specific words that are commonly used in the domain. For instance, if an individual attends a university and studies a particular academic subject (e.g., psychology, education, nursing), then he/she must learn the technical words that are common in the field. Learning these words, along with high-frequency ones, can provide L2 learners with a stable level of comprehension in different texts (Dang et al., 2022).

A well-known word list for academic purposes is the *Academic Word List* (Coxhead, 2000), which consists of approximately 3.5 million running words from more than 400 academic texts in 28 academic fields (e.g., arts, commerce, law, science). It includes a total of 570 word families (WFs) that were not included in the GSL. An example of the specific word lists is the *Nursing Academic Word List* (NAWL) (Yang, 2015). Yang extracted 676 WFs from the *Nursing Research Academic Corpus* consisting of 1,006,934 running words from 253 nursing research articles in 21 nursing specialties (e.g., cardiology, oncology, rehabilitation).

In general, different word lists have been created for various purposes. However, what these lists have in common is that they were extracted from corpora that matched the learning purpose. According to Browne (2014, p. 2), "Any word frequency lists that are generated from a corpus will be a direct reflection of the texts in that corpus." Therefore, in order to create CWLs, we built an original *English Cartoon Corpus* (ECC) and extracted the commonly used words.

Rationale for this study

In general, there are four aspects to consider when creating word lists: (1) how the corpus is prepared; (2) how the words are counted; (3) the criteria for selecting appropriate words (e.g., frequency, range, dispersion); and (4) how to determine the validity of the list. Previous studies have shown that there are two main ways to prepare a corpus from which to extract using existing large corpora (e.g., CEC, COCA, BNC) (Brezina & Gablasova, 2015; Browne, 2013; Dang et al., 2017; Gardner & Davies, 2014) or building a new corpus (Coxhead, 2000; Yang, 2015). Regarding the first concern, we built an original corpus, the ECC.

As for how the words are counted, an inevitable debate when creating word lists is what unit to adopt for counting words (Cobb & Laufer, 2021). Previous studies have mainly employed the lemma (Browne, 2013; Brezina & Gablasova, 2015; Gardner & Davies, 2014) or the WF (Coxhead, 2000; Dang et al., 2017; Yang, 2015). A WF includes a headword (e.g., play), its inflected forms (e.g., plays, played, playing), and its derived forms (e.g., replay, player, playful). Conversely, a lemma only consists of a headword and its inflected forms. This difference means that a WF groups more words together than a lemma. Meanwhile, although a flemma (i.e., family lemma) is similar to a lemma, it does not distinguish between parts of speech (POS) when grouping word forms that share the same spelling. For example, *back* includes various usages as a noun (e.g., My *back* hurts.), an adjective (e.g., Use the *back* door.), an adverb (e.g., Come *back* soon.), and a verb (e.g., Many people *backed* my proposal.). Here, the lemma counts *back* as four words, whereas the flemma counts it as one word.

Proponents of the WF have argued that knowing the headword (e.g., aware) makes it easier to predict the meaning of the derivatives (e.g., awareness, unaware), whereas those for the lemma have argued that there are many derivatives (e.g., invaluable, particle) that cannot be predicted based on the meaning of the headword (e.g., value, part) (Brezina & Gablasova, 2015). Additionally, the WF can lead to an overestimation of the learners' vocabulary size (Gardner & Davies, 2014). For instance, it is assumed that a learner who knows the headword (e.g., resist) also knows the inflected words (e.g., resists, resisted, resisting). However, this does not necessarily guarantee that the learner knows all the derivatives included in a WF (e.g., resistive, resistible, resistivity), particularly those whose meanings are not readily predictable from the meaning and form of the base word. Although the lemma appears to be a more appropriate unit than the WF, this study adopted the flemma, which is in line with the NGSL 1.2. As discussed later, we found some instances of incorrect misclassification of POS information in the lexical analysis tool. Thus, it would be more reliable to create the CWLs without distinguishing between POS when counting word forms (see Data analysis).

When creating word lists, frequency is the best criterion for extracting the words most used in a particular domain. Coxhead (2000) listed the words that appeared 100 or more times within the academic written corpus of approximately 3.5 million words (i.e., 28.57 per million words, PMW), while Yang (2015) included the words with a frequency of 33 or more times in 1,006,943 running words (i.e., 32.78 PMW). Similarly, Dang et al. (2017) set the criterion at 350 or more times in the academic spoken corpus of 13,029,661 running words (i.e., 26.9 PMW). Thus, we set the criterion at 30 or more PMW.

Along with frequency, range could also be a criterion. The greatest drawback of range is that a particular word is counted if it occurs only once in a corpus. Therefore, the dispersion (Juilland's D) value (0.01 to 1.0) becomes important, which shows how evenly a particular word appears in the entire corpus. In this case, the closer the value to 1.0, the more evenly

distributed the word. In line with the previous study (Dang et al., 2017), this study set the dispersion threshold at 0.6 or higher.

Finally, when determining the validity of the created word lists, the most important quality is the lexical coverage, the percentage of listed words in given texts (Cobb & Laufer, 2021). In the present study, we determined how well the CWLs achieved the lexical coverage in the new cartoon scripts. The research questions are as follows:

RQ1. How many words beyond high-frequency words occur frequently and widely in cartoons?

RQ2. What percentage of the words in the new cartoon scripts can the CWLs cover?

Methods

Creating the English cartoon corpus (ECC)

In this study, we created the ECC from the ground up. Specifically, we selected cartoons streamed from Netflix because the *Language Reactor* extension allowed us to collect more running words by simply copying and pasting the English captions into a text file, and because the recommended viewing age and genre information (e.g., comedy, horror, action) were listed for each cartoon. For cartoons originally produced in languages other than English, we selected the English-dubbed audio and English captions on Netflix. To verify the correspondence between the English captions and the spoken language, we randomly selected one minute from the beginning, middle, and latter part of each episode, resulting in three minutes of audio per episode, and confirmed that the captions corresponded to the language actually spoken by the characters and narrators. It should be noted that the pasted English captions often contained words not spoken in the actual scenes (e.g., [laughs], [birds chirping], [dramatic theme playing]). These words were excluded from the text files to maintain consistency in the dataset. The captions were used solely for corpus construction and linguistic analysis in this study. The original captions are not reproduced or redistributed in this article; only aggregated lexical statistics and the resulting word lists are reported.

Netflix also provides recommended age ratings (i.e., All, 7+, 10+, 13+, 16+, 18+) for each cartoon. Since the warning labels for the 10+ and 13+ cartoons are similar, and the 16+ and 18+ cartoons are suitable for adults, this study categorized them as follows: (1) all-age cartoons for toddlers; (2) 7+ cartoons for young children; (3) 10+ and 13+ cartoons for older children; and (4) 16+ and 18+ cartoons for adults. Based on this classification, four English cartoon subcorpora were created: the ECC for Toddlers, Young Children, Older Children, and Adults, respectively. In previous studies, the total number of running words collected by researchers ranged from approximately one million (Yang, 2015) to 3.5 million words (Coxhead, 2000). We ensured that each ECC contained at least one million words.

Data analysis

In this study, the ideal lexical analysis tool should include the following three functions, in order of priority: (1) to analyze as many running words as possible (at least eight million words); (2) to calculate the dispersion values; and (3) to lemmatize English words. To the best of our knowledge, however, there is no existing lexical analysis tool that meets all of these requirements. Meanwhile, the tool that meets the first and second conditions (but not the third) is *#LancsBox* (Brezina et al., 2021). Thus, we selected this tool for data analysis.

Overall, the use of *#LancsBox* revealed that its POS classification was not entirely accurate. For example, in the ECC for Toddlers, *delicious* appeared 397 times, of which 75 were classified as nouns and the remainder were classified as adjectives. However, *delicious* is not

used as a noun in English. POS tagging was not sufficiently reliable for lemma-based counting; therefore, the flemma was adopted for creating the CWLs.

To address RQ1, the lemmas that met the criteria of a frequency of 30 or more PMW and a dispersion value of 0.6 or higher were extracted. However, as illustrated by the case of *delicious*, an incorrect POS classification can lead to the total frequency being distributed. For instance, in the ECC for Toddlers, which consists of 2,434,903 running words (Table 1), the frequency threshold of 30 PMW corresponded to a minimum of 74 occurrences ($[30 \times 2,434,903] / 1,000,000$). If word *X* is classified as occurring 40 times as a noun and 40 times as an adjective, then it would not meet the minimum frequency threshold of 74 occurrences and be excluded as a word list candidate. To overcome this issue, we automatically aggregated the frequencies of identical spellings across POS categories by using the UNIQUE and SUMIF functions in Excel. Although frequencies were aggregated across POS to create flemmas, the dispersion value of the most frequent POS was retained to ensure that the most frequent usage of each spelling was sufficiently widespread in the corpus. In this case, if words occurred at least 74 times, and if the dispersion value of the most frequent POS was 0.6 or higher, then the words were included as candidates. For example, if word *Y* has both noun and verb usages, and appears in the corpus 30 times as a verb (dispersion = 0.30) and 44 times as a noun (dispersion = 0.60), then it was considered a candidate. The remaining candidates were then manually checked by the authors against the predefined exclusion criteria. Words that did not meet these criteria were excluded to finalize the list. Specifically, the excluded items included proper nouns (e.g., America, Teotihuacan, George), numbers (e.g., one, two, 3), ordinal numbers (e.g., 1st, 2nd, third), cardinal numbers (e.g., hundred, thousand, million), marginal words (e.g., grr, ohh, umm), greetings (e.g., hello, bye, ta-ta), days of the week (e.g., Sunday, Monday), months (e.g., January, November), letters (e.g., a, b, c), holidays (e.g., Christmas, Halloween), acronyms (e.g., NATO, DNA), and colloquial expressions (e.g., wanna, gimme, 'sup). This process was repeated for each ECC, with the goal of producing the four CWLs.

Overall, the word list that best fit the purpose of this study was the *TV/Movie Corpus* (TVMC) (<https://www.wordfrequency.info/samples.asp>). As for RQ2, the lexical coverage provided by the CWLs and the TVMC was compared by using the scripts of other cartoons that were not included in the ECC. The TVMC comprises the top 5,050 lemmas ranked by frequency. We converted these lemmas into 4,177 flemmas by using the UNIQUE and SUMIF functions. For lexical coverage comparisons of other cartoons, the same number of flemmas was extracted from the TVMC as from the CWLs. For example, if the CWLs contained 100 flemmas, the top 100 flemmas were extracted from the TVMC.

Results and Discussion

CWLs for toddlers, young children, older children, and adults

Table 1 presents the descriptive statistics for each subcorpus. The analysis of the ECC for Toddlers showed that the corpus initially contained 31,501 flemmas before applying the selection criteria, of which 372 were identified as final candidates based on the frequency and dispersion criteria for the CWL for Toddlers. Then, we double-checked each word against the predefined exclusion criteria to determine whether the remaining words were appropriate for the list. Finally, 271 flemmas were included in the CWL for Toddlers. Following the same process, we included 266 and 137 flemmas in the CWL for Young Children and Older Children, respectively. Because many flemmas (e.g., asleep, magical, worm) were shared across the three lists, the three individual CWLs contained 674 flemma entries ($271 + 266 + 137$), whereas the combined list contained 452 distinct flemmas (see the Supplementary Material). However, the majority of the CWL for Adults contained insult slang (e.g., jerk, asshole, bastard), sexual slang

(e.g., dick, tit, boner), aggressive slang (e.g., fuck, bitch, shit), violent expressions (e.g., kill, choke), and derogatory expressions (e.g., trash, whore). Since this CWL was not appropriate from an L2 teaching perspective, we did not conduct further analyses of it, include it in the Supplementary Material, or make it publicly available.

We summarize the characteristics of the words in each CWL. For instance, the words included only in the CWL for Toddlers were dominated by food-related words (e.g., berry, coconut, bake, chew) and animal-related words (e.g., butterfly, dinosaur, dragon, beak, fluffy). Additionally, the list included educational themes designed to stimulate intellectual curiosity, such as space (e.g., astronaut, comet), light (e.g., dazzle, glitter), plants (e.g., leaf, lily), and nature (e.g., pond, rainbow).

Meanwhile, the words in both the CWL for Toddlers and Young Children contained words related to animals and food (e.g., puppy, banana) as well as friendship (e.g., buddy, teamwork). However, the CWL for Young Children was about adventure such as fighting/battle (e.g., conquer, commander), weapons (e.g., arrow, shield), rivals/enemies (e.g., kidnap, villain), friendship (e.g., backup, sidekick, squad), and success (e.g., congratulations, glory). As for the common words in both the CWL for Young Children and Older Children, they shared more stimulating expressions (e.g., betray, curse, nightmare) than the CWL for Toddlers.

Finally, the words listed in the CWL for Older Children included more violent words (e.g., banish, chaos, invasion) and insulting ones (e.g., nerd, weirdo) than those in the CWL for Toddlers and Young Children. These words were influenced by the fact that the cartoons for Older Children dealt with fantasy and supernatural phenomena (e.g., angel, heaven).

Table 1

Data for Each English Cartoon Subcorpus

	Toddlers	Young Children	Older Children	Adults	Total
Tokens	2,434,903	2,160,410	1,226,478	2,195,125	8,016,916
Lemmas	37,640	50,017	34,906	51,439	104,693
Flemmas	31,501	42,209	29,038	44,418	90,268
Works	44	42	26	37	149
Episodes	1,377	1,067	600	712	3,756
Broadcast Year	2020.44	2019.31	2019.31	2020.31	2019.97
<i>SD</i>	2.66	3.76	4.28	2.58	3.26
Range	2009–2024	2005–2024	2006–2024	2013–2024	2005–2024

Note. The numbers of lemmas and flemmas represent all items in each corpus before applying the predefined exclusion criteria for the CWLs.

Lexical coverage

To determine the validity of the CWLs, the lexical coverage of the CWLs and the TVMC were compared. We extracted the top 452 flemmas from the 4,177 flemmas in the TVMC. Among the 452 flemmas from the CWLs and the TVMC, 208 flemmas (46.0%) were common to both lists.

To compare lexical coverage, three other types of scripts (i.e., cartoons for toddlers, young children, and older children) were collected that were not included in the ECC. For each script, one episode per cartoon was randomly selected, for a total of 10 episodes over 10 cartoons. Table 2 presents the number of flemmas and tokens in each corpus occupied by the NGSL 1.2,

including 208 common flemmas in the CWLs and the TVMC (*Common*), 244 flemmas in the CWLs (*CWLs*), 244 flemmas in the TVMC (*TVMC*), and all other words, except those in the aforementioned lists (*Off the Lists*). *Others* include words that met the predefined exclusion criteria and were therefore excluded from the CWLs, such as proper nouns, marginal words, numbers, and days of the week (see Data analysis). Additionally, the 2,809 flemmas in the NGSL 1.2 were divided into five levels in order to match the 452 flemmas in the CWLs: (1) the top 1,001 flemmas in the NGSL 1.2 (Nos. 1–1,001, Level 1); (2) the next 452 flemmas (Nos. 1,002–1,453, Level 2); (3) the next 452 flemmas (Nos. 1,454–1,905, Level 3); (4) the next 452 flemmas (Nos. 1,906–2,357, Level 4); and (5) the final 452 flemmas (Nos. 2,358–2,809, Level 5). This made it possible to determine which CWLs were of equal value to which levels in the NGSL 1.2. The reason for not further dividing Level 1 was that the level included approximately 1,000 flemmas, which could be considered the most frequent 1,000 flemmas. Note that *Others* were not included in the calculation of the lexical coverage.

Table 2 also shows that there are 40,888 tokens in the 10 episodes of cartoons for toddlers. Since 2,707 tokens were included as *Others*, the remaining 38,181 tokens were used to calculate the lexical coverage. The NGSL 1.2 covered more than 90% (91.01%), while the lexical coverage of 452 flemmas in the CWLs was 4.15%, which (together with the NGSL 1.2) exceeded 95% (95.16%). Conversely, the other 452 flemmas in the TVMC covered 2.41%. This indicated that the CWLs provided higher lexical coverage than the TVMC for cartoons for toddlers. Interestingly, the lexical coverage of the 452 flemmas in the CWLs was higher than that of Level 2 (4.15% vs. 2.84%). This suggested that the 452 flemmas in the CWLs provided greater lexical coverage than the 452 flemmas in Level 2.

Table 2
Data for the Three Types of Cartoons

	Cartoons for Toddlers		Cartoons for Young Children		Cartoons for Older Children	
Lists	Flemmas	Tokens	Flemmas	Tokens	Flemmas	Tokens
Level 1 (1,001)	696	31,914 (83.59)	829	38,251 (85.61)	881	51,850 (86.44)
Level 2 (452)	213	1,085 (86.43)	260	1,194 (88.28)	293	1,459 (88.87)
Level 3 (452)	152	612 (88.03)	201	898 (90.29)	239	1,084 (90.68)
Level 4 (452)	129	528 (89.41)	197	696 (91.85)	197	740 (91.92)
Level 5 (452)	115	611 (91.01)	168	509 (92.99)	190	753 (93.17)
<i>Common</i> (208)	115	795 (93.10)	153	724 (94.61)	156	786 (94.48)
CWLs (244)	115	789 (95.16)	125	367 (95.43)	123	427 (95.19)
TVMC (244)	42	127 (95.50)	82	175 (95.82)	98	363 (95.80)
<i>Off the Lists</i>	694	1,720 (100.0)	1,152	1,868 (100.0)	1,430	2,520 (100.0)
Total	2,271	38,181	3,167	44,682	3,607	59,982
<i>Others</i>	378	2,707	730	3,272	688	3,972

Note. The values in parentheses in the tokens indicate the cumulative lexical coverage (%).

Analysis of the 10 episodes of cartoons for young children also revealed that the lexical coverage in the NGSL 1.2 exceeded 90% (92.99%). Specifically, the lexical coverage in the 452 flemmas in the CWLs was found to be 2.44%, while that for the combination of the NGSL 1.2 and the CWLs exceeded 95% (95.43%). Moreover, the lexical coverage of the 452 flemmas in the CWLs was comparable to Level 2 (2.44% vs. 2.67%) and higher than Level 3 (2.44% vs. 2.01%). Similarly, analysis of the 10 episodes of cartoons for older children revealed that the NGSL 1.2 covered over 90% (93.17%), while the lexical coverage of the 452 flemmas in the CWLs was 2.02%. This indicated that the sum of the NGSL 1.2 and the CWLs exceeded 95% (95.19%). In contrast, the TVMC did not reach 2.0% lexical coverage (1.92%). Meanwhile, the lexical coverage of the 452 flemmas in the CWLs was comparable to Level 2 (2.02% vs. 2.43%) and higher than Level 3 (2.02% vs. 1.81%). As for the lexical coverage in the cartoons for older children, it was the lowest (2.02%), compared to those for toddlers and young children (4.15%, 2.44%). This was possibly due to the fact that the number of flemmas in the CWL for Older Children (137 flemmas) was smaller than those of the CWLs for Toddlers and Young Children (271, 266 flemmas, respectively).

Conclusion

This study created CWLs for L2 learning based on extensive viewing of cartoons. The main findings were as follows: (1) the NGSL 1.2 alone provided more than 90% lexical coverage; (2) the 452 flemmas in the CWLs provided higher lexical coverage than the 452 flemmas in the TVMC; (3) a total of 3,261 flemmas in both the NGSL 1.2 and the CWLs exceeded 95% lexical coverage; and (4) the lexical coverage of the 452 flemmas in the CWLs was higher than or comparable to those of Level 2 and higher than Level 3. The implication of the findings is that learning the CWLs after the first 1,000 flemmas in the NGSL is effective for increasing lexical coverage of cartoons. Because receptive knowledge of the words used in cartoons facilitates comprehension during cartoon viewing, the CWLs may help learners benefit more from extensive viewing as a source of authentic audiovisual input. This finding underscores the importance of domain-specific word lists for improving comprehension in audiovisual input.

EFL learners may benefit from regular exposure to a significant amount of comprehensible and authentic audiovisual input to support their L2 development. Streaming services (e.g., Netflix, Amazon Prime, Disney+) include extensive viewing materials that can help solve the problem of limited exposure to English outside the classroom. In fact, the total number of episodes in the ECC constructed in this study was 3,756.

Cartoons can support incidental learning of high-frequency words. In fact, toddlers and young children can acquire high-frequency words by watching cartoons (Alexiou, 2015; Arifani, 2020). Thus, cartoons may serve as a useful first step toward continuous L2 learning, even for adult learners. Cartoons, especially for toddlers, frequently use a *here-and-now* approach, which emphasizes the use of language related to present situations, objects, events, or actions. Even when viewers watch a cartoon without spoken English (e.g., Tom and Jerry), they can predict how the story will unfold.

There are several limitations that should be noted, the first of which is related to the lexical analysis tool. The ECC was constructed from captions representing spoken English. Hence, the lexical analysis software could not analyze the abbreviations of colloquial expressions. For example, the present participle forms (e.g., doing, breaking) were often abbreviated in the captions (e.g., doin', breakin'). However, the software did not correctly lemmatize these words and instead treated them as distinct word forms. If these words had been properly lemmatized back to their original form, then more verbs might have been included in the CWLs.

Second, the adoption of the flemma as the counting unit has one important limitation. Because the CWLs were created without distinguishing between different POS, the Supplementary Material file does not indicate which POS is the most frequent for each flemma. Therefore, when learning the CWLs, learners should be aware that each flemma may represent multiple POS. For example, *bow* is most commonly used as a verb (e.g., The student *bowed* politely before leaving the room.), but it is also used as a noun (e.g., The sailors stood on the *bow* of the ship.). Consequently, learning flemmas requires understanding multiple meanings associated with different POS, which may take more time than learning lemmas.

Finally, although the pedagogical implications of the CWLs are promising, this study evaluated their usefulness only through lexical coverage analyses rather than learner-based evidence or experimental validation. Future research should examine whether learning the CWLs improves learners' comprehension during cartoon viewing and supports vocabulary acquisition.

Note

The Cartoon Word Lists can be downloaded as the [Supplementary Material here](#).

References

- Alexiou, T. (2015). Vocabulary uptake from Peppa Pig: A case study of preschool EFL learners in Greece. In C. Gitsaki, & T. Alexiou (Eds.), *Current Issues in Second/Foreign Language Teaching and Teacher Development: Research and Practice* (pp. 285–301). Cambridge Scholars Publishing.
- Arifani, Y. (2020). Cartoon video-assisted learning: An investigation into the acquisition of EFL children's incidental vocabulary. *Computer-Assisted Language Learning Electronic Journal*, 21(2), 17–31. <https://callej.org/index.php/journal/article/view/295>
- Browne, C. (2013). The new general service list: Celebrating 60 years of vocabulary learning. *The Language Teacher*, 37, 13–16. https://jalt-publications.org/tlt/issues/2013-07_37.4
- Browne, C. (2014). A new general service list: The better mousetrap we have been looking for? *Vocabulary Learning and Instruction*, 3, 1–10. <https://doi.org/10.7820/vli.v03.2.browne>
- Browne, C., Culligan, B., & Phillips, J. (2024). *New general service list project*. <https://www.newgeneralservicelist.com>
- Brezina, V., & Gablasova, D. (2015). Is there a core general vocabulary? Introducing the new general service list. *Applied Linguistics*, 36, 1–22. <https://doi.org/10.1093/applin/amt018>
- Brezina, V., Weill-Tessier, P., & McEnery, A. (2021). #LancsBox v. 6.x. [software package]. <https://lancsbox.lancs.ac.uk>
- Cobb, T., & Laufer, B. (2021). The nuclear word family list: A list of the most frequent family members, including base and affixed words. *Language Learning*, 71, 834–871. <https://doi.org/10.1111/lang.12452>
- Coxhead, A. (2000). A new academic word list. *TESOL Quarterly*, 34, 213–238. <https://doi.org/10.2307/3587951>
- Dang, T. N. Y., Coxhead, A., & Webb, S. (2017). The academic spoken word list. *Language Learning*, 67, 959–997. <https://doi.org/10.1111/lang.12253>
- Dang, T. N. Y., Webb, S., & Coxhead, A. (2022). Evaluating lists of high-frequency words: Teachers' and learners' perspectives. *Language Teaching Research*, 26, 617–641. <https://doi.org/10.1177/1362168820911189>
- Gardner, D., & Davies, M. (2014). A new academic vocabulary list. *Applied Linguistics*, 35, 305–327. <https://doi.org/10.1093/applin/amt015>
- Hu, M., & Nation, P. (2000). Unknown vocabulary density and reading comprehension. *Reading in a Foreign Language*, 13, 403–430. <https://doi.org/10.64152/10125/66973>

- Kanayama, K. (2024). Effects of extensive viewing of cartoons on L2 listening improvement: How much audiovisual input is needed? *Annual Review of English Language Education in Japan*, 35, 1–16. https://doi.org/10.20581/arele.35.0_1
- Kanayama, K. (2026). Do captions affect listening? Long-term effects on listening proficiency and comprehension. *International Journal of Applied Linguistics*, Early View. <https://doi.org/10.1111/ijal.70235>
- Kurokawa, S., Hein, A. M., & Uchihara, T. (2025). Incidental vocabulary acquisition through captioned viewing: A meta-analysis. *Language Learning*, 75(4), 939–987. <https://doi.org/10.1111/lang.12697>
- Masrai, A. (2020). Can L2 phonological vocabulary knowledge and listening comprehension be developed through extensive movie viewing? The case of Arab EFL learners. *International Journal of Listening*, 34, 54–69. <https://doi.org/10.1080/10904018.2019.1582346>
- Mayer, R. E., & Fiorella, L. (2022). *The Cambridge handbook of multimedia learning*. Cambridge University Press.
- Rodgers, M. P., & Webb, S. (2017). The effects of captions on EFL learners' comprehension of English-language television programs. *CALICO Journal*, 34, 20–38. <https://doi.org/10.1558/cj.29522>
- Teng, M. F. (2021). *Language learning through captioned videos*. Routledge.
- Webb, S. (2015). Extensive viewing: Language learning through watching television. In D. Nunan & J. C. Richards (Eds.), *Language learning beyond the classroom* (pp. 159–168). Routledge.
- Webb, S., Uchihara, T., & Yanagisawa, A. (2023). How effective is second language incidental vocabulary learning? A meta-analysis. *Language Teaching*, 56, 161–180. <https://doi.org/10.1017/S0261444822000507>
- West, M. (1953). *A general service list of English words*. Longman.
- Yang, M. N. (2015). A nursing academic word list. *English for Specific Purposes*, 37, 27–38. <https://doi.org/10.1016/j.esp.2014.05.003>

Kohei Kanayama is an Associate Professor at Hokkaido University of Education. He holds a Master's degree in Education from Hokkaido University of Education, Japan. His research focuses on vocabulary acquisition and listening instruction through extensive viewing. He is particularly interested in corpus-based approaches to language learning and the development of materials for learning English through viewing cartoons.

Akira Iwata is a Professor at Hokkai-Gakuen University. He previously taught at five public high schools for 21 years and at Hokkaido Musashi Women's Junior College for nine years. He holds a PhD from Tokyo Gakugei University, Japan, and a Master's degree in Education (TESOL) from the University of Wollongong, Australia. His research focuses on EFL classroom practices, particularly vocabulary acquisition and reading within EFL pedagogy.