

## Most Frequently Used Vocabulary in Selected English Drama Movie Scripts

Sawitri Suwanaroa<sup>1\*</sup> and Sutarat Polerk<sup>2</sup>

<sup>1</sup>Faculty of Business Administration and Liberal Arts, Rajamangala University of Technology Lanna Tak, Thailand

<sup>2</sup>Faculty of Business Administration and Liberal Arts, Rajamangala University of Technology Lanna Tak, Thailand

Corresponding Author: Sawitri Suwanaroa, E-mail: [sawitri\\_75@yahoo.com](mailto:sawitri_75@yahoo.com)

### ARTICLE INFORMATION

**Received:** November 02, 2020

**Accepted:** December 25, 2020

**Volume:** 3

**Issue:** 12

**DOI:** 10.32996/ijllt.2020.3.12.19

### KEYWORDS

Language, content words, English drama movie scripts, part of speech, vocabulary, descriptive research

### ABSTRACT

Understanding words in motion pictures is a rewarding pathway for audience of cultural-based media and entertainment. The study aimed to investigate the most frequently used vocabulary in selected English drama movie scripts. It sought to find out the most frequently occurred part of speech and to probe the frequency of occurrences of each content words in the movie scripts. Five English drama movie scripts were chosen purposively as sources of the corpora. Applying the Antconc program software developed by Anthony, a total of 108 content words were noted and categorized for corpora analysis. There were 38 verbs, 28 adverbs, 24 nouns and 18 adjectives noted as commonly used vocabulary from the movie scripts. Among the verb types, have (366), can (276) and know (232) emerged as the most frequently occurred action words with a total frequency of 3,142 occurrences. In the case of adverbs, here (237), out (178), and about (163) came out as the most frequently used words with a total occurrence of 1,598. Nouns, on the other hand, have 1,574 occurrences with time (134), tree (133), and going (127) as the most frequently used name words. Adjective was found to be the least frequently occurred content word with 752 occurrences to include right (146), old (54), and only (53) as the top three frequently used noun modifiers. In conclusion, verb outnumbered the other three content words due to drama's nature of involving the characters into action; and 'have' came out as the most frequently used verb because drama scenes imply possessions and obligations.

### 1. Introduction

For several years, 'many countries include English as a second language in their school syllabus and children start learning English at a young age' (Baker, 2008). This has made one-fifth of world citizen is able to either communicate or understand English. Besides, 53 countries entitle English as an official language. This is an inevitable evidence of the most widespread language in the world.

Mass media provides significant chances for foreign language speakers to use authentic ways to listen and communicate in English. Baker (2008) claims that television is one of the best mass media for information contribution such as political, cultural, cognitive and educational. Generally, television is a platform of personal's development. Kosonen (2008) described the popularity of television creates the English language process more obvious, persuasive and emotional. Information chances of television flows the foreign language audience to listen and get used to with authentic ways of communication (Schmitt, 2000). It also reflects life of society, unique features of national mentality, national stereotype, that is why television can be considered as a source of information (Kirkpatrick, 2010). Regarding to Baker (2009), television programs are the entertainment source of audiences' knowledge, experience, emotions, and impressions. These essential points create the integrated culture and language through mass media that use as a device to attract audiences from different countries. Additionally, many strategies including words used in drama movies are used by film industries in order to convey audiences (Pennycook, 2010). Then, the spread of English language and the rapid growth of technology creates the use of subtitle in television and movies.

From American Awards to Goldsmith Prize and Billboards Top 100, English language has evidently been a mode of textual composition. With exposure to English language, learners will certainly be comprehensible to extensive world of cultural understanding together with the greater world. "Everyone knows that Hollywood is in the United States, and that the biggest television and music industries in the world are based there. As it mentioned above, TV and movies are a great way to practice your English once you start learning. What's more, knowing English opens up thousands of movies, television shows, and games for your enjoyment" (Foerch and Kasper, 1983). Therefore, the researchers attempted to study the frequently used vocabularies for drama movie scripts 'Hidden Figures, Passengers, The Best Offer, The Terminal, Forest Gump' as the sample of data sources (Wohl, 2008).

Words that used in movies depend on genre of the movie. If the movie is an action or others, words must be specific of each genre. Using word by according to the type of movie will generate realistic feeling to viewer (Katzner, 2002). Therefore, you can learn a lot of different vocabularies by watching same types of movies. Alternatively, learner can learn English Language through movies and can apply to use their English knowledge on their own way.

### **1.1 Research questions**

This study seeks to find answers to the following questions:

1. What content words that are frequently used in drama movies?
2. Which type of content words are found in the drama movies that have the highest frequency?

### **2. Literature review**

In this study, the researchers focus on the pragmatic features of interjections used in communication to find out the meaning and contexts that each interjection was used in. The samples of the study are movie scripts of 6 films from 3 genre, action, drama, and thriller. The researchers categorized interjection from the sample into 16 types of interjections based on the films and contexts. Resulted in the communicative content of interjections is vaguer than the text books would predict or covered the entire information. Also, one interjection can fall into more than one category depends on the context it was used in.

In "Vocabulary size text coverage" Nation and Waring (2008) did a study about how vocabulary size and word list which second language learner need while comparing to the vocabulary size of native speaker.

This research was purposely undertaken as to establish an ultimate goal for vocabulary learning.

On the ground that frequency information involves the learning output as a result of learners' effort, it is necessary for curriculum design and learning goals to incorporate vocabulary frequency as a significant element.

By saying this, instead of constructing an extensive vocabulary list, the curriculum designer should be better deal with some vocabulary reference lists that particularly assist in language course design. Therefore, only demanding vocabulary are being considered as well as resulting in its suitability for different classes.

The availability of powerful computers that is definitely large corpora currently make the development of such lists a much easier task than it was when Boers (2013) and their colleagues manually counted 18,000,000 running words.

The contribution of word frequency list is therefore a result of judgment-based construction of corpora with the use of well-established criteria.

In summary of the study, this research underpinned the deciding factors which are likely to be developed: Representativeness, Frequency and range, Word families, Idioms and set expressions, Range of information. This is in accordance with the study of Nation and McCarthy (1997) which investigated the boundary between general service vocabulary and special purposes vocabulary.

To illustrate the experiment, this paper adopted the use of some criteria: frequency, text coverage, and range as to investigate dividing line between a general service vocabulary and a special purposes vocabulary.

Whilst a general service vocabulary could contain up to 2000-word level, Crystal (1997) suggested that learners in contrary, who are requisite only those properly managed corpora are likely to demonstrate well in learning vocabulary efficacy in addition to its primary aim to establish guiding plan to the design of course for specific purposes.

As a result, this study stated that the 2000-word general service vocabulary is sufficient for prospective learners at specific purposes level. When completing the acquisition of 2000-word general service vocabulary, learners can greatly extend text coverage by studying a new frequency word list.

Frequently used words are words that constantly used in daily communication. They can cover most of the contents in any contexts in English. It can be said that by learning or acknowledge these words, most ESL learners can instantly understand what they read or hear. Although frequently used words or general service list may suggest around 2000 words, it may not cover 100% of the contents and after these 2000 words may require knowledge of more specialized vocabulary. Thus, the learners can improve their abilities of automatic recognition which lead to fluency in English communication by studying the type of words which included content word and function word. By collecting information from the films, the researchers will analyst the population of content word from the films and display the result of content word population in the films.

In summary, based on the previous study of Boers (2013), the text of 18,000,000 running tokens were manually counted in order to output the result of vocabulary frequency.

Another research conducted by Nation and Waring (2008) revealed that corpora can effectively improve learners' vocabularies pool. Moreover, Nation and McCarthy (1997) carried out a study on general word and technical word which purposively to investigate the vocabularies from the drama movies which has an influential toward ESL learners.

**3. Research methodology**

The samples were scripts of five movies with the same genres. The scripts were from "Hidden Figures, Passengers, The Best Offer, The Terminal and Forest Gump". The researcher firstly had watched five movies to get understanding of the plot through stories. The length of each script was between 8,500 and 12,000 words. These words are retrieved from entire scripts of movies. The research instruments used in the study were a corpus and a concordance the Antconc Program.

In the study, five movie scripts were collected to compile a corpus. Then, the Antconc Program was used to count words and build a word frequency list. The listed words were sorted in a frequency order, and words with less than 20 occurrences in the corpus were deleted. Next words were sorted in alphabetical order. Only content words for example nouns, verbs, adjectives and adverbs were selected to make a new list of high frequency content words. The researcher created a table in order to show the high frequency of content word.

**4. Results and discussion**

The results of the study showed that 108 content words with at least 20 occurrences in the corpus appeared in five drama movies scripts. Table 1 shows the result of a high frequency list of content words. There were 21 words in the list that their frequency was over 100 occurrences in the frequency rank: have (366), can (276), here (237), know (237), go (230), get (218), out (178), like (177), above (163), how (150), right (146), come (142), time (134), tree (133), well (131), going (127), think (115), back (113), see (109), tell (103) and never (100).

In content words high frequency list, there were top six words in list based on Table 1: have, can, here, know, go and get. "Have" is one of those words with a hundred applicants. It means to hold or maintain as a possession, privilege, or entitlement. "Can" mean to be able to through acquired knowledge or skill. "Here" means using when gesturing to indicate the place intended or draw attention to someone or something that has just arrived. "Know" has to be familiar with or have experience and know also mean "understand" or "knowledge" for example someone who is in the know has information about something. "Go" means leave a place, especially in order to travel to somewhere else. "Get" means to catch or grab and also be used to take about direction for example if someone to receive or be given something and to go somewhere or bring back someone or something. The speech pattern of saying "Get" incessantly don't get, mean anything.

**Table 1:** The result of high frequency list of content word

No	Word	F	No	word	F	No	word	F	No	Word	F
1	Have	366	31	Life	62	61	Away	35	91	Live	27
2	Can	276	32	Home	57	62	Call	35	92	Woman	27
3	Here	237	33	House	57	63	Wrong	35	93	After	26
4	Know	232	34	Always	56	64	Sorry	34	94	Last	26

5	Go	230	35	Help	56	65	Stay	34	95	Feel	25
6	Get	218	36	Old	54	66	Best	33	96	Idea	25
7	Out	178	37	Only	53	67	Believe	33	97	Late	25
8	Like	177	38	Space	53	68	Wide	32	98	Read	25
9	Above	163	39	Keep	52	69	Talk	33	99	Show	25
10	How	150	40	Good	50	70	Else	32	100	Afraid	24
11	Right	146	41	Little	48	71	Flight	32	101	Laugh	24
12	Come	142	42	Clair	47	72	Ship	32	102	Great	24
13	Time	134	43	Bad	47	73	Table	32	103	Body	23
14	Tree	133	44	Earn	47	74	War	32	104	Remember	22
15	Well	131	45	Find	47	75	Beautiful	32	105	Try	22
16	Going	127	46	Understand	45	76	Same	31	106	Problem	21
17	Think	115	47	Wait	45	77	Friend	31	107	Hear	20
18	Back	113	48	Again	45	78	School	31	108	Really	20
19	See	109	49	Shrimp	42	79	Real	31			
20	Tell	103	50	God	41	80	Ark	31			
21	Never	100	51	Around	40	81	Looking	31			
22	Still	95	52	Ever	40	82	Mean	30			
23	Need	87	53	Lieutenant	40	83	Door	29			
24	Take	87	54	Before	39	84	Leave	29			
25	Down	78	55	Thing	39	85	Room	29			
26	Make	76	56	Give	38	86	Today	29			
27	Quickly	74	57	Night	37	87	Country	28			
28	Thank	74	58	Eat	36	88	Earth	28			
29	Look	70	59	Maybe	36	89	Fine	28			
30	People	76	60	North	36	90	Through	28			

**Note:** No = number of word  
F = frequency

After classifying words to make a noun list, function words and words specific to particular movies were deleted. The resulting lists of nouns were shown in Table 2. It was found that there were 32 nouns in the list. Words in the frequency rank were time (134), tree (133), going (127), back (113), people (66), life (62), home (58) and house (57). In other 25 words included space (47) clair (47), shrimp (42), god (41), lieutenant (40), thing (39), night (37), north (36), flight (32), ship (32), table (32), war (32), friend (31), school (31), looking (30), door (29), room (29), today (29), country (28), earth (28), woman (27), idea (25), stupid (25), body (23) and problem (20) respectively.

**Table 2:** The result of a noun high frequency list

No	Noun	F	No	Noun	F	No	Noun	F	No	Noun	F
1	Time	134	11	Shrimp	42	21	Friend	31	31	Body	23
2	Tree	133	12	God	41	22	School	31	32	Problem	20
3	Going	127	13	Lieutenant	40	23	Looking	30			
4	Back	113	14	Thing	39	24	Door	29			
5	People	76	15	Night	37	25	Room	29			
6	Life	62	16	North	36	26	Today	29			
7	Home	58	17	Flight	32	27	Country	28			
8	House	57	18	Ship	32	28	Earth	28			
9	Space	47	19	Table	32	29	Woman	27			
10	Clair	47	20	War	32	30	Idea	25			

**Note:** No = number of word

F = frequency

After classifying words to make a verb list, function words and words specific to particular movies were deleted. The resulting lists of verbs were shown in Table 3. It was found that there were 38 verbs in the list. Words in the frequency rank were have (366), can (276), know (232), go (230), get (218), like (177), come (142), think (115) and see (109). The other 28 words included tell (103), need (87), take (87), make (76), thank (74), look (70), help (56), keep (52), earn (47), find (47), understand (45), wait (45), give (38), eat (36), call (35), stay (34), believe (33), talk (33), ark (30), mean (30), leave (29), live (27), fell (25), read (25), show (25), laugh (24), remember (22), try (22) and hear (20) respectively.

**Table 3:** The result of verb high frequency list

No	Verb	F	No	Verb	F	No	Verb	F	No	Verb	F
1	Have	366	11	Need	87	21	Wait	45	31	Live	27
2	Can	276	12	Take	87	22	Give	38	32	Fell	25
3	Know	232	13	Make	76	23	Eat	36	33	Read	25
4	Go	230	14	Thank	74	24	Call	35	34	Show	25
5	Get	218	15	Look	70	25	Stay	34	35	Laugh	24
6	Like	177	16	Help	56	26	Believe	33	36	Remember	22

7	Come	142	17	Keep	52	27	Talk	33	37	Try	22
8	Think	115	18	Earn	47	28	Ark	30	38	Hear	20
9	See	109	19	Find	47	29	Mean	30	<b>Note:</b> No = number of word F = frequency		
10	Tell	103	20	Understand	45	30	Leave	29			

After classifying words to make an adverb list, function words and words specific to particular movies were deleted. The resulting lists of adverbs were shown in Table 4. It was found that there were 20 adverbs in the list. Words in the frequency rank were here (237), out (178), about (163), how (150), now (141), well (131), never (100), still (95) and down (78). The other 18 words included quickly (74), always (56), again (40), around (40), ever (40), before (39), maybe (36), away (35), best (33) else (32), through (28), after (26) and really (20) respectively.

**Table 4:** The result of an adverb high frequency list

No	Adverb	F	No	Adverb	F
1	Here	237	11	Again	40
2	Out	178	12	Around	40
3	Above	163	13	Ever	40
4	How	150	14	Before	39
5	Well	131	15	Maybe	36
6	Never	100	16	Away	35
7	Still	95	17	Else	32
8	Down	78	18	Through	28
9	Quickly	74	19	After	26
10	Always	56	20	Really	20

**Note:** No = number of word

F = frequency

After classifying words to make an adjective list, function words and words specific to particular movies were deleted. The resulting lists of adjectives were shown in Table 5. It was found that there were 18 adjectives in the list. Words in the frequency rank were right, old (54), only (53), good (50), little (48), bad (47), wrong (35), Sorry (34), best (33), wide (32), beautiful (31), same (31) and real (31). The other 18 words included fine (28), through (28), last (26), afraid (24), great (24) and fine (23) respectively.

**Table 5:** The result of an adjective high frequency list

No	Adjective	F	No	Adjective	F
1	Right	146	11	Beautiful	31

2	Old	54	12	Same	31
3	Only	53	13	Real	31
4	Good	50	14	Fine	28
5	Little	48	15	Last	26
6	Bad	47	16	Late	25
7	Wrong	35	17	Afraid	24
8	Sorry	34	18	Great	24
9	Best	33	19		
10	Wide	32	20		

**Note:** No = number of word

F = frequency

In brief, words appearing in drama movies are mostly used in dairy life due to drama movies instituted real life. Normally, drama movies could not be complicated words which used high frequency words in other genres of movies and dependent on the associations of the characters, age and roles. Specifically, in drama movies, there are not only words that appearing, they consist of idioms, slangs, jokes or others to show the acting or emotional expression of the characters. Furthermore, the high frequency words in drama movies are interesting to study but not only in the movies and it should be also in the other forms and means of entertainment. High frequency words in drama movies could be analysed in many English media, conversation dairy life, games and books.

**5. Discussion of the Study**

The objective of this study is to analyse frequently used words from drama movie scripts and to investigate type of content words found in drama movie scripts that has high frequency. Improving English proficiency by watching movies is one of a very effective way that is also popular among people all over the world. (Wohl, 2008). Furthermore, watching movies to improve English listening skill is a simple and entertaining method. Therefore, it inspires the researchers to choose scripts from drama movies as samples for the study since this kind of movies tend to be based on actual events. Vocabularies found in the movie scripts are mostly simple and understandable for those who would like to learn more vocabularies from the film scripts, also it can be easily adapted to use in real life situations. By learning more vocabularies from the movie scripts, learners can improve their English proficiency in terms of listening and communication method almost efficiency. Therefore, English learners can learn about frequently used vocabulary from films scripts to have a better basic of English proficiency, such as speaking and listening skills.

Data used in this study is movie scripts. The researchers choose 5 drama movie scripts including, Hidden Figures, Passengers, The Best Offer, The Terminal, and Forest Gump as samples in this study. The most frequently used vocabulary was counted in those five movies by Antconc Program, Anthony. Based on the database, there are 5 movie scripts which are saved as text files in folder before counting with Antconc program to find the most frequently used vocabulary in the film scripts used in this study to create table of frequently used vocabulary. Vocabularies will be sorted according to their frequency from lowest to highest. The researchers choose only vocabularies that appear at least 20 times in the movie scripts. Also, those vocabularies must be placed under four categories namely Noun, Verb, Adverb and Adjective according to the part of speech in English grammar.

The result in this study of high frequency words was completed. The high frequency content words were counted, there are 108 words. Words resulting were classified to their parts of speech to set 4 lists of high frequency words which content words as verb are the most commonly used. High frequency list of verbs, there were top five words in list: have (366), can (276), here (237), know (232) and go (230). "Have" is one of those words with a hundred applicants. It means to hold or maintain as a possession, privilege, or entitlement. "Can" means to be able to through acquired knowledge or skill. "Here" means using when gesturing to indicate the place intended or draw attention to someone or something that has just arrived. "Know" has

to be familiar with or have experience and know also means “understand” or “knowledge” for example someone who is in the know has information about something. “Go” means leave a place, especially in order to travel to somewhere else. And followed by noun, adverb and adjective respectively.

In conclusion, according to the first research question “What content words that frequently used in drama movies?” can be described that the most found content word according to the table 1 which divided into noun, verb, adverb, and adjective for a total of 117 words. To answer the second research question, “Which type of content words found in the drama movies that has highest frequency?” the answer to this question is verb, since it has the highest frequency which is 38 words, next to verb is noun which has 33 words, adverb 28 words, and that last one is adjective which has 18 words respectively.

The result of the study shows the words that mostly used in drama movie which tend to relate to learners’ daily life. Also, there are many words can be used in the classroom to expand the vocabularies pool of learners which lead to the improvement of their communication skills. This study focuses on frequently used words in the movies to expand the vocabularies pool which learners can integrate into their English knowledge.

### 7. Limitation and suggestion for future research

Results of this study should be interpreted critically and carefully because of the following reasons. First reason is the number of movie scripts can be critically selected based on the content of the movie. Secondly, it should not be only specific on English movie, it could be possible to investigate on Thai-English code-mixing movie or drama. Also, the following part is a suggestion for those who interested in further investigation in the contents related to this study. From the analysis of 4 types of content word, it can be concluded that verb is the most found type. Consequently, the further study can be investigated on how verbs were used in the sentence based on movie script. This will be a beneficial information for those would like to learn English via watching movies. Also, learners can improve their knowledge of idioms, slangs, drama movies sub-titles or any other movies to improve their English. Not only vocabularies, idioms, and other elements, but learners can also learn English-Thai code-mixing from the movie scripts. Moreover, those who interested can learn cultures and beliefs from the films or can practice their listening and speaking skill from the movies to improve their proficiency.

### References

- [1] Baker, W. (2008). A critical examination of ELT in Thailand: the role of cultural awareness. *RELC Journal*, 39(1), 131-146.
- [2] Boers, F. (2013). Cognitive linguistic approaches to teaching vocabulary: *Assessment and integration*. *Language Teaching*, 46(02), 208-224. <https://doi.org/10.1017/S0261444811000450>
- [3] Crystal, D. (1997). *English as a global language*. Cambridge University Press.
- [4] Foerch, C. & G. Kasper. (1983). *Plans and strategies in foreign language communication*. In Færch, C. & G. Kasper (eds.), 20-60.
- [5] Katzner, K. (2002). *The Languages of the World*. Routledge.
- [6] Kirkpatrick, A. (2010). *English as a lingua franca in ASEAN*. Hong Kong University Press.
- [7] Kosonen, K. (2008). Literacy in Local Languages in Thailand: Language Maintenance in a Globalised World. *International Journal of Bilingual Education and Bilingualism*, 11(2), 150-184.
- [8] Schmitt, N., & McCarthy, M. (1997). *Vocabulary: Description, Acquisition and Pedagogy* (pp. 6-19). Cambridge University Press.
- [9] Nation, P. and Waring, R. (2008). *Vocabulary size text coverage*. Cambridge University Press.
- [10] Pennycook, A. (2010). *Popular cultures, popular languages, and global identities*. In N. Coupland (Eds.), *the handbook of language and globalization*. Blackwell.
- [11] Schmitt, N. (2000). *Vocabulary in language teaching*. Cambridge University Press.
- [12] Spears, R. A. (1993). *NTC's dictionary of phrasal verbs and other idiomatic verbal phrases*. National Textbook Company.
- [13] Spears, R. A. (2000). *NTC's dictionary of American slang and colloquial expressions*. NTC Publication Group.
- [14] Wohl, M. (2008). *The language of film*. Ken Stone. [http://www.kenstone.net/fcp\\_homepage/language\\_of\\_film.html](http://www.kenstone.net/fcp_homepage/language_of_film.html)