

# Corpus Analysis of the Police Language in the American TV Series

## *CSI: NY*

\*Anchan Premjai<sup>1</sup>

1999orkey@ygmail.com

Jonathan Rante Carreon<sup>2</sup>

carreonjrc@gmail.com

Areeluck Harnmontree<sup>3</sup>

areeluckharn@gmail.com

Huachiew Chalermprakiet University<sup>1,2,3</sup>

\*corresponding author

### Abstract

The field of English for Specific Purposes has continuously broadened its scope especially among the members of the expanding circle of English speaker (Kachru, 1992; 2006). This research paper examines the content of the American TV Series *CSI: NY* to identify the words with highest absolute frequencies keywords or the words with highest relative frequencies (e.g. Scott, 1997, 2000), and N-grams or combination of words (e.g. Bednarek, 2011) used by police officers in the TV series while doing their jobs. Relative frequencies were obtained by comparing the data against the British National Corpus (BNC) and computing their log-likelihood values, and the N-grams were identified from the concordance of each keyword using the software Antconc (Anthony, 2017). The keywords and n-grams were then iteratively categorized into themes and compared against the actual functions of police officers. With a high cut-off log-likelihood value of 100.00, the findings showed 71 keywords divided into four categories: (1) informal spoken language use (N=26, 36.62%), (2) addressing oneself, other people or things (N=24, 33.80%), (3) crime scenes (N=16, 22.54%), and (4) interviewing suspects, witnesses, or other persons (N=5, 7.04%). Pedagogically, the findings have strong implications since the words with the highest absolute frequencies, and the n-gram are indispensable in preparing ESP (listening and speaking) courses for police students and officers who are members of the expanding circle of English speakers (Kachru, 2006).

**Keywords:** English for Specific Purposes, keyword analysis, N-gram, police language

### 1. Introduction

From the early beginnings of English for Specific Purposes (ESP; Hutchinson & Waters, 1987) until the present time, a huge amount of work has already been done. Specifically, courses are taught for the various sectors of business (Business English, English for Traders, English for Investors, and others). Moreover, English courses tailored for the medical and allied medical professions, lawyers, aviation industry, tourism, hotels and restaurants are also quite common. The contents of these courses are aimed mainly at developing the four language skills of listening, speaking, reading and writing as applied in a specific profession. Nevertheless, one more profession which is concerned with security, and peace and order that may eventually use English in their career in the ASEAN region is the police profession.

Police officer, “a social-security service career”, is a profession that includes constant interaction and communication with both the locals and foreign population of a country. In Thailand, airport immigration officials and tourist police who screen in coming of foreigners and look at the security needs of foreigners entering Thailand, autonomously, have high English-speaking proficiencies. In contrast, the local police officers have relatively lower

English proficiency levels. These police officers may find it difficult to function effectively if they are unable to use a language understandable to foreigners.

## 2. Related research

In western countries, Fenner, Gudjonsson & Clare (2002) studied about people's understanding of police cautions in England and Wales. This was later expanded to explore the complexity of police cautions and the extent of understanding these cautions (Eastwood & Snook, 2010 & Eastwood, Snook & Chaulk, 2012). These suggest that police language can be misunderstood and pose challenges to police officers in making themselves clear and understandable when communicating with concerned people. Studying the understanding of police cautions is valuable for both police officers and people. However, real life police officers are not limited to use only cautions when they are on duty. Expressions for informing, asking, reporting, and others are also used. This observation suggests that there is a need to explore police language in purposes other than cautions.

In Thailand, many research studies have examined the use of English language in the police profession. Tansrisawat (1991), Promrat (1998), Meemark (2002), and Tipmontree (2007) aimed to analyze tourist police officers' needs and problem in English communication skills. Khamkaew (2009) studied English communication of rural tourist police officers and the problems they encountered while providing help at the tourist counter service area. A survey was also conducted to obtain Thais and foreigners' viewpoints on tourist police's role and language skill abilities after the official opening of the AEC in 2015 and found that Thais and foreigners expect the tourist police will have high English proficiency (Wichasin & Dounghummes, 2014).

These past research studies, however, focused mainly on police language communication problems, people's perspective on police language skills and police language competencies and undermined other crucial issue in language learning, such as the source of materials used in developing English courses for teaching English language skills for Thai police officers. Specifically, this research study examined the data to identify words with high absolute frequencies and relative frequencies, and word combinations as used in the police profession.

Examining language at the word level may already shed light on the important vocabularies used in the police profession. However, people communicate not only by using words, but also by combining and recycling words. One way to present combinations of words is by using the N-gram model. The N-gram model focuses on presenting N (number) of items (words) that are put together as a sequence of text or speech (Broder, Glassman, Manasse, & Zweig, 1997). Bednarek (2008) applied to use N-gram model to categorize and examine language presented in a movie, and continued exploring about spoken language in the TV series *Gilmore Girls* in 2011 and 2013.

From the same token, this research study applies N-gram model to identify and classify word combinations in police language that are suitable and useful because most police language learners aim to learn "language in action (Jones, 2012, p.27)". The police officers mainly use English to communicate to each other, such as foreign victim, suspect, witness, or even police in other police stations. Therefore, having a grasp of what these language combinations are provide police officers studying English a more straightforward set of "language in action".

### **3. Methodology**

#### **3.1 Data and data collection**

The police TV series of choice is the *CSI: NY*. While there is a large number of movies and TV series in the world depicting police activities, *CSI: NY* is chosen to be the data source of the study because firstly, the movie was inspired by real police stories. It was assumed that characters portray their roles as if it were real. This is crucial because this study is interested in naturally-occurring word combinations. Secondly, the word combinations in the data will be compared against the word combinations from a benchmark. The benchmark or comparator corpus is the British National Corpus (BNC), which is a corpus of general language use, composed of about 100 million words. Thus, to eliminate any potential discrepancies such as poor choice of words, grammatical inconsistencies and other language issues, the police TV series played by native English speakers is chosen. Finally, the researcher would like to analyze word combinations taken from police real life, so *CSI: NY* is drama that film genre that depicts realistic police characters and roles.

The scripts of all the seventeen episodes of *CSI: NY Season 9* were downloaded from [www.subscene.com](http://www.subscene.com). Before analysis, ordinary people's dialogs were removed to make sure that the data is mainly composed of police dialogs.

#### **3.2 Data analysis**

The data analysis consists of two parts. Firstly, a collection of police words in *CSI: NY Season 9* is run using free software used in corpus analysis called AntConc (Bednarek, 2008) to find absolute frequencies. Secondly, the words with absolute frequencies of at least 20 were compared with the British National Corpus (BNC) to explore the differences between police and general language use. The outputs in this stage are words with relative frequencies or also known as the keywords (e.g. Scott, 1997, 2000). Put another way, any words with a log-likelihood of greater than 10 were considered keywords. These keywords within their local co-text were then categorized using an iterative process of identifying themes. Four themes were identified:

- Words relating to addressing oneself, other people or things.
- Words relating to formal spoken language use.
- Words relating to interviewing suspects, witnesses, or other persons.
- Words relating to crime scenes.

Next, the identified keywords were iteratively thematized using the co-text produced by each keyword following four steps: (1) divide the keywords into sets of 20 keywords (e.g. Carreon & Watson Todd, 2013), (2) examine the concordance lines of each keyword (Bednarek, 2011), (3) categorize the keywords based on police purposes (Baton Rouge Police department, 2017), and (4) quantify the themes. Finally, the keywords related to police purposes were examined for N-grams.

### **4. Results**

#### **4.1 Absolute and relative frequencies**

The initial analysis of absolute frequencies yielded 47,836 tokens and 4,731 word types. To identify the keywords, the authors chose the words that have absolute frequencies of 20 and higher. The top 5 words with the highest absolute frequencies are shown in Table 1 below.

**Table 1** Top 5 words with highest absolute frequencies

No.	Tokens/Words	Frequencies
1	The	1885
2	You	1482
3	I	1106
4	To	1097
5	a	1071

Some initial implications can be drawn from Table 1. The high frequency of *the* reflects the use long running sentences which allows the interlocutors to show definiteness (e.g. *the suspect, the site, and the evidence*) and also to show coherence in their ideas. The high frequencies of the first person pronouns *you* and *I* suggest that the data is conversational in nature. The pronouns *you* and *I* are used in English conversation to address the first person speaker or the second person interlocutor. It is difficult to reach stronger conclusions, however, as absolute frequencies reflect general language use as well as the concerns of the particular texts under investigation (Carreon & Watson Todd, 2013).

Thus, to identify the words that more directly characterize the data, relative frequencies were examined. As mentioned earlier, these were calculated by comparing the absolute frequencies against frequencies in the BNC using log-likelihood (see Rayson & Garside, 2000 for a detailed discussion of log-likelihood). Words with log-likelihood values of 100.00 and higher are deemed keywords of the *CSI: NY Season 9* TV series. There is no clear existing literature as to what log-likelihood values should be taken to show that the words are “key” or salient to the data. It is widely argued, however, that the highest the log-likelihood values of the words, the more that they reflect the concerns of that particular data (Carreon, Watson Todd & Knox, 2011). The same point of reasoning has been embraced by research study.

There are 79 words with log-likelihood values 100.00 and higher after comparing with the BNC. These keywords are iteratively thematized using their respective local co-text as guide. The 71 keywords with log-likelihood values of 100 and higher are given in Table 2.

**Table 2** Keywords with the highest log-likelihood values (1-70)

Keywords List							
No.	Tokens/Words	*F	**LL	No.	Tokens/Words	*F	*LL
1	s ('s)	1047	8243.90	41	cop	25	186.30
2	you	1482	2221.80	42	guys	31	184.76
3	re ('re)	224	1922.36	43	which	31	180.36
4	don (don't)	181	1559.05	44	tell	88	177.33
5	m ('m)	234	1444.75	45	blood	56	175.14
6	ve ('ve)	103	1210.50	46	let	78	158.70
7	uh	75	800.74	47	jo	25	156.51
8	I	1106	783.45	48	why	106	156.07
9	hey	96	717.42	49	jason	25	147.84
10	vic	72	616.11	50	huh	21	142.68
11	mac	68	562.40	51	looks	52	142.53
12	d	107	480.66	52	felipe	20	140.50
13	gun	72	404.39	53	in	596	137.63
14	guy	72	386.42	54	hopkins	20	134.57
15	what	383	385.26	55	out	227	133.31

16	your	275	377.76	56	someone	60	131.39
17	brooks	43	374.23	57	killed	44	129.92
18	me	269	369.68	58	kill	35	129.86
19	benny	47	367.10	59	our	140	129.30
20	look	66	358.95	60	detective	25	127.19
21	how	128	356.15	61	all	245	125.60
22	got	217	353.12	62	shot	42	125.17
23	okay	95	352.64	63	no	242	120.98
24	christine	49	343.48	64	want	100	119.14
25	sid	39	333.30	65	ashley	21	115.01
26	right	208	323.08	66	guess	28	114.08
27	leonard	42	291.16	67	like	177	112.27
28	know	218	263.95	68	get	134	110.75
29	we	415	257.05	69	name	64	108.73
30	here	159	251.95	70	tommy	20	104.11
31	him	237	228.38	71	hell	31	100.49
32	he	604	225.37				
33	murder	54	221.53				
34	justin	26	218.17				
35	killer	36	215.63				
36	yeah	162	210.80				
37	milner	21	203.52				
38	adam	42	200.25				
39	found	117	196.20				
40	kid	35	196.11				

\*F = Frequency, \*\*LL = Log-likelihood

From Table 2, words with high relative frequencies differ from words with high absolute frequencies and they reflect more accurately the main content of the data in question. These keywords were categorized into four themes: (1) words relating to addressing oneself, other people or things; (2) words relating to informal spoken language language use; (3) words relating to interviewing suspects, witnesses, or other persons; and (4) words relating to crime scene. The reliability of this categorization was rated excellent (Cohen's kappa=0.953). The highest-ranked keywords for each theme with example of use are given in Table 3. For reading convenience and illustration, only the first 24 keywords are shown in this section since all the categories can already be found up this number. The remaining keywords can be found in the Appendix.

**Table 3** Keywords and themes

No.	Keywords	*F	**LL	Concordance	Themes
1	you	1482	2221.80	1. But we can bring the scene to <b>you</b> . 2. The phone log shows <b>you</b> called the vic 17 times...	(1) Addressing oneself, other people or things
2	I	1106	783.45	1. <b>I</b> called it in. 2. <b>I</b> can't remember the simplest things.	(1) Addressing oneself, other people or things
3	mac	68	562.40	1. <b>Mac</b> , it's Mr.Lewis.	(1) Addressing

				2. <b>Mac</b> Taylor in the lounge with the Glock.	oneself, other people or things
4	your	275	377.76	1. We got <b>your</b> blood, we got <b>your</b> prints. 2. It's someone you invited onto <b>your</b> boat...	(1) Addressing oneself, other people or things
5	brooks	43	374.23	1. Video recordings of Leonard <b>Brooks</b> ' prison therapy sessions. 2. <b>Brooks</b> used basic chemistry to kill Jimmy Clark.	(1) Addressing oneself, other people or things
6	me	269	369.68	1. You brought <b>me</b> out here. 2. Do you want to tell <b>me</b> why you're really here?	(1) Addressing oneself, other people or things
7	benny	47	367.10	1. <b>Benny</b> 's blood was all over it. 2. Raymond shocked <b>Benny</b> with a car battery...	(1) Addressing oneself, other people or things
8	christine	49	343.48	1. We found <b>Christine</b> 's phone at a jewelry store... 2. Any sign of <b>Christine</b> or Shawn Boyd?	(1) Addressing oneself, other people or things
9	s ('s)	1047	8243.90	1. Our vic's a male. 2. There's a bank drop-off two blocks from here.	(2) Informal spoken language use
10	re ('re)	224	1922.36	1. You're under arrest for assault and resisting arrest. 2. We're searching for a missing young girl.	(2) Informal spoken language use
11	don (don't)	181	1559.05	1. Janitors <b>don't</b> get here till 10.00... 2. <b>Don't</b> move!	(2) Informal spoken language use
12	m ('m)	234	1444.75	1. I'm Detective Lovato with the NYPD. 2. I'm gonna need that patient's name.	(2) Informal spoken language use
13	ve ('ve)	103	1210.50	1. We've just been looking in the wrong place. 2. He could've tossed the gun on the roof, maybe.	(2) Informal spoken language use
14	uh	75	800.74	1. And, <b>uh</b> , maybe somebody can tell us her real name. 2. I'm, <b>uh</b> ...I'm pretty sure that our guy's...	(2) Informal spoken language use
15	hey	96	717.42	1. <b>Hey</b> , Flack, I got your text. 2. <b>Hey!</b> Calm down!	(2) Informal spoken language use
16	vic	72	616.11	1. Our <b>vic</b> is Ellen White, 19 years old. 2. Our <b>vic</b> was driving the van.	(2) Informal spoken language use
17	would	107	480.66	1. You'd rather leave a man die in the	(2) Informal

	('d)			street... 2. ...how they'd take the news of his murder...	spoken language use
18	guy	72	386.42	1. We have a dead <b>guy</b> . 2. This <b>guy</b> knows that you saw his face.	(2) Informal spoken language use
19	look	66	358.95	1. <b>Look</b> familiar? 2. He got a <b>look</b> at the guy who went over the fence.	(2) Informal spoken language use
20	got	217	353.12	1. It's look like <b>got</b> a few good prints there. 2. He's <b>got</b> a gun. 3. Just <b>got</b> a hit on Interpol. 4. We <b>got</b> a hit off some prints on the latex gloves...	(2) Informal spoken language use
21	okay	95	352.64	1. <b>Okay</b> . 2. <b>Okay</b> ? 3. Is he <b>okay</b> ?	(2) Informal spoken language use
22	gun	72	404.39	1. Guy who took a shot at you dumped the <b>gun</b> . 2. That <b>gun</b> has been around the block a few times. 3. ...Nathan hid his <b>gun</b> in the attic... 4. Looks like he's putting the <b>gun</b> in your hands. 5. ...you brought a <b>gun</b> to a public place...	(3) Crime scenes
23	what	383	385.26	1. <b>What</b> can you tell us about these? 2. <b>What</b> 's up with this guy? 3. Tell me <b>what</b> you did to her! 4. <b>What</b> about the rest of the video? 5. <b>What</b> happened after you abandoned your vehicle?	(4) Interviewing suspect, witness, or person
24	how	128	356.15	1. Tell me <b>how</b> a gun registered in your name... 2. <b>How</b> about murder? 3. <b>How</b> 'd she afford(s) a dress like that? 4. <b>How</b> did your blood end up on her dress? 5. <b>How</b> 'd you light the fires, Leonard?	(4) Interviewing suspect, witness, or person

The first two themes in Table 3 reflect the data which is spoken in nature. Thus, we might expect similar keywords to the first and second themes in other communications concerning the other episodes of the *CSI: NY Season 9 TV Series* and other TV series of similar genre. It can also be safely argued that since these two themes are characteristic features of spoken language, they can also be found in TV series of other genres and in conversational exchanges. The third and fourth themes are directly related to the functions of a police officer as categorized by the Baton Rouge Police Department (2017) in the U.S. For a complete list of the iterative thematic analysis, please see the Appendix section.

**Table 4** Themes, frequencies and their percentages

Themes	Total	%
Informal spoken language use	26	36.62
Addressing oneself, other people or things	24	33.80
Crime scenes	16	22.54
Interviewing suspects, witnesses, or other persons	5	7.04
<b>TOTAL</b>	<b>71</b>	<b>100.00</b>

**Table 4** shows that words relating to *informal spoken language use* (N=26, 36.62%) and words relating to *addressing oneself, other people or things* (N=24, 33.80%) dominate the data followed by words relating to *crime scenes* (N=16, 22.54%) and words relating to *interviewing suspects, witnesses, or other persons* (N=5, 7.04%). From these figures, it is not surprising that the keywords in the first two themes are more highly frequent than those that are directly related to their jobs since police officers are not always responding to emergencies in crime scenes or interviewing suspects or witnesses throughout the time they are on duty. Nevertheless, these keywords still shed light on the nature of spoken language police officers use while they are on-duty and have pedagogical implications, particularly in preparing material for teaching. For example, the concordance of the keywords in the third theme can be an authentic source of examples for teaching police officers the skill of interviewing, such as asking and answering questions. The keywords in the theme crime scenes are further examined for word combinations they produce.

#### 4.2 N-grams

There were 16 keywords under the theme words relating to *crime scenes* (N=16, 22.54%). These keywords produce N-gram ranging from 2-grams to 6-grams. For example, the keyword *shot* (F=42; LL=125.17), produce 2-grams (e.g. *shot at*), 3-grams (e.g. *shot at me, take a shot*), 4-grams (e.g. *a shot at me, a shot at Hopkins*), 5-grams (e.g. *taking a shot at Scarlet, take a shot at Hopkins*), and 6-grams (e.g. *heard a shot at that point*).

The 2-gram *shot at* is a phrasal verb following the combination verb + preposition. The first 3-gram *shot at me* follows the combination verb + preposition + noun/pronoun. The second representative 3-gram *take a shot* follows the combination verb + article + noun. The representative 4-gram *a shot at me* follows the combination article/no article (plurals) + verb + preposition + noun/pronoun. The 5-gram *take/taking a shot at Scarlet* follows the combination verb + article + preposition + noun. Figure 1 is a snapshot of the keyword *shot*.

to forcibly enter the precinct. The mope took a shot at him. That's what happened. You saw him  
 . That's what happened. You saw him take a shot at Hopkins? No. You hear the shot? No. So  
 . So how do you know that Brown took a shot at Hopkins? Because Kevin said so, and that  
 . So you think the vic really did take a shot at Hopkins? I think he thinks the vic took  
 Hopkins? I think he thinks the vic took a shot at Hopkins. But the truth is, no one knows  
 it. He had a gun; the kid took a shot at me. You called it in? Yeah, I called  
 hands! Kevin, you all right?! Yeah, he took a shot at me. Dispatcher: 198, come in to Central. 1  
 lie. I am not lying. That kid took a shot at me. I called a 13. Stay there. Don't  
 you. That kid, he pointed a gun and he shot at me, and I will go to my grave  
 Emerson for this, and I buy her taking a shot at Scarlet and hanging White, but I don't  
 doubled back to find Hopkins. Had you heard a shot at that point? No, but that doesn't mean  
 and killed with your gun. Guy who took a shot at you dumped the gun. Two kids found it.

**Figure 1** Snapshot of *shot*



## 5. Discussion and Conclusion

First, let us state that most analysis of keywords consider log-likelihood values that are much lower than the values used in this study (e.g. Carreon & Watson Todd, 2013). Setting a higher ceiling of cut-off value produces 'key' keywords (Scott, 1997) that are indispensable in starting basic English training course. At the beginning of this paper, it was mentioned that the ultimate goal of this paper is to identify keywords that characterize the *CSI: NY Season 9 TV Series* and whether these keywords match any major police function and therefore useful as source of training material for developing the English speaking skill of police officers in EFL countries like Thailand.

The main findings revealed 71 keywords with log-likelihood of 100.00 and higher. While the percentage of keywords directly related to the police profession is only around 21 keywords or about 29.58% of 71 keywords, these keywords confirm the categories of police functions proposed by the Baton Rouge Police Department (2017) in the U.S. The bigger portion of 50 keywords or about 70.42% of the keywords depicts a picture of police headquarters with police officers who are on-duty but not attending to any emergency situation or not talking to a witness, suspect or criminal. Put it another way, these police officers are doing general office chores like chatting with colleagues or doing basic office tasks. While these keywords are not directly related to police functions, they shed some light on how police offices communicate with their colleagues and therefore can be useful especially for beginner police language learners in EFL countries like Thailand.

As for the n-grams, they depict the kind of language structures police officers may use while they are doing their functions. These are indispensable in learning police language since they provide pattern of language use (Bednarek, 2012). In other words, word combinations in the shape of n-grams also provide frames of language production and help ESL/EFL learners decode individual words or interpret whole meaning (Zhao, 2009). Many ESL/EFL classes teach word combinations because they are formed with the grammar of the language (Krishnamurthy, 2002).

One main limitation of this research is the sole focus on keywords. In reality, police officers who are planning to take basic English language courses related to the field may have limited English language background, and may be forced to study the more difficult keywords or jargons instead of starting from the foundation level. Thus, there is a need to examine also the combinations of function words in the *CSI: NY* series to identify, for example, how sentences are formed, the basic word combinations used and others. We hope that this corpus-based investigation of police keywords furthered some knowledge and understanding on possible source of data for preparing English language training for police officers especially for members of the expanding circle of English speaker (Kachru, 2006).

## References

- Anthony, L. (2014). AntConc (Version 3.4.3) [Computer software]. Tokyo: Waseda University. Accessed 9 July. <http://www.laurenceanthony.net/>.
- Baton Rouge Police Department. (2017, July 30). Retrieved from Baton Rouge Police Department: [brgov.com/dept/brpd/pdf/police\\_duties.pdf](http://brgov.com/dept/brpd/pdf/police_duties.pdf)
- Bednarek, M. (2011). The language of fictional television: A case study of the 'dramedy' *Gilmore Girls*. *English Text Construction*, 4(1), 54-84.
- Bednarek, M. (2012). "Get us the hell out of here": Key words and trigrams in fictional television series. *International Journal of Corpus Linguistics*, 17(1), 35-63.
- Bednarek, M. A. (2008). 'What the hell is wrong with you?' A corpus perspective on evaluation and emotion in contemporary American pop culture.

- Broder, A. Z., Glassman, S. C., Manasse, M. S., & Zweig, G. (1997). Syntactic clustering of the web. *29 (8)*, 1157–1166.
- Carreon, J. R., & Todd, R. W. (2013). Conflicts between prioritizing medical care and profit-making for a Thai hospital. *International Journal of Language Studies*, *7(1)* pp.117-138.
- Carreon, J. R., Todd, R. W., & Knox, J. (2013). Medical tourism communication of a Thai private hospital website. *Journal of Applied Linguistics and Professional Practice*, *8(2)*, 165-185.
- Eastwood, J., & Snook, B. (2012). The effect of listenability factors on the comprehension of police cautions. *Law and Human Behavior*, *36(3)*, 177.
- Eastwood, J., Snook, B., & Chaulk, S. (2010). Measuring reading complexity and listening comprehension of Canadian police cautions. *Criminal Justice and Behavior*, *37(4)*, 453-471.
- Fenner, S., Gudjonsson, G., & Clare, I. (2002). Understanding of the current police caution (England and Wales) among suspects in police detention. *Journal of Community & Applied Social Psychology*, *12(2)*, 83-93.
- Hutchinson, T. &. (1987). ESP A learning centred approach. *Civilacademy*. Cambridge: Cambridge University Press., *7(1)*, 27-46.
- Jones, R. H. (2012). *Discourse analysis*. New York.: Abingdon/New York.
- Kachru, Y., & Nelson, C. L. (2006). *World Englishes in Asian Contexts* (Vol. 1). Hong Kong University Press.
- Khamkaew, S. (2009). *Needs and problems in English listening and speaking skills: A case study of metropolitan police officers at counter service as Chana Songhram police station*. Bangkok: Srinakharinwirot University.
- Kirkpatrick, A. (2008). English as the official working language of the Association of Southeast Asian Nations (ASEAN): Features and strategies. *English Today*, *24(02)*, 27-34.
- Krishnamurthy, R. (2002). Language as chunks, not words. In *Jalt 2002 Conference Proceedings* (pp. 288-294). Tokyo: JALT Publication.
- Meemark, M. (2002). *An analysis of needs and problems of English for Tourist Police*. Bangkok: Mahidol University.
- Promrat, N. (1998). *A Study of English Vocabulary Competency and Problems of the Tourist Police in the Lower Central Region of Thailand*. Pitsanulok: Naresuan University.
- Rayson, P. &. (2000, October). Comparing corpora using frequency profiling. In *Proceedings of the workshop on Comparing Corpora*. Association for Computational Linguistics, (pp. 1-6).
- Scott, M. (1997). PC analysis of key words—and key key words. *System*, *25(2)*, 233-245.
- Scott, M. (2000). Focusing on the text and its key words. *Rethinking language pedagogy from a corpus perspective*, *2*, 103-122.
- Tansrisawat, S. (1991). *An opinion study of the concerned-persons in tourism industry towards the operation of tourist police in Pattaya*. Bangkok: Thammasart University.
- Television, J. B. (Director). (2012-2013). *CSI:NY Season 9* [Motion Picture].
- Tipmontree, S. (2007). *The Use and the Problems of English and Intercultural Communication Skills of Thai Tourist Police Officers (Doctoral dissertation)*. Bangkok: University of the Thai Chamber of Commerce.
- Wichasin, P., & Doungphummes, N. (2014). Roles of Thai tourist police in Asean Economic Community. *Journal of Management Cases*, *Vol.16(3)*, 40-58.
- Zhao, Y. (2009). *An empirical of L2 learners' use of lexical chunks and language production*.

## Appendix

No.	Keywords	*F	**LL	Concordance	Themes
25	leonard	42	291.16	1. Got the suspect, <b>Leonard</b> Brooks. 2. <b>Leonard's</b> in the wind.	(1) Addressing oneself, other people or things
26	we	415	257.05	1. <b>We</b> found a bag of pills in his backpack. 2. Okay, <b>we</b> found a print on the jeweler's safe.	(1) Addressing oneself, other people or things
27	him	237	228.38	1. You saw <b>him</b> at the scene. 2. Somebody could have driven the car after <b>him</b> .	(1) Addressing oneself, other people or things
28	he	604	225.37	1. Well, <b>he</b> has to be in the crowd. 2. <b>He</b> ran his prints through AFIS, got nothing back.	(1) Addressing oneself, other people or things
29	justin	26	218.17	1. He found <b>Justin</b> dead on the floor. 2. Check <b>Justin's</b> GPS records.	(1) Addressing oneself, other people or things
30	milner	21	203.52	1. Keith <b>Milner</b> , the only suspect in the disappearance of Tommy... 2. What can you tell me about Keith <b>Milner's</b> murder?	(1) Addressing oneself, other people or things
31	adam	42	200.25	1. <b>Adam</b> and Hawkes don't have detective shields. 2. <b>Adam</b> are on foot in the alley heading northbound toward 122.	(1) Addressing oneself, other people or things
32	jo	25	156.51	1. I'm <b>Jo</b> Danville, crime lab. 2. The champagne bottle that <b>Jo</b> found.	(1) Addressing oneself, other people or things
33	jason	25	147.84	1. Hi, I'm <b>Jason</b> Riley. 2. <b>Jason</b> trusted you, and you betrayed that trust.	(1) Addressing oneself, other people or things
34	felipe	20	140.50	1. <b>Felipe</b> dragged Tortucci out of the van... 2. <b>Felipe</b> Zacharias is in the wind.	(1) Addressing oneself, other people or things
35	hopkins	20	134.57	1. We got to get <b>Hopkins</b> and Jensen out of here. 2. <b>Hopkins</b> is lying about the gun.	(1) Addressing oneself, other people or things
36	someone	60	131.39	1. He had to have told <b>someone</b> about that address. 2. Looks like he's got <b>someone</b> following in his footsteps.	(1) Addressing oneself, other people or things

37	our	140	129.30	1. I want to give him <b>our</b> address. 2. ...this footage from one of <b>our</b> Department surveillance camera...	(1) Addressing oneself, other people or things
38	detective	25	127.19	1. I'm <b>detective</b> Messer with the NYPD. 2. ...I am a first-grade <b>detective</b> ...	(1) Addressing oneself, other people or things
39	ashley	21	115.01	1. ...where was <b>Ashley</b> before she was at the park? 2. When did you last see <b>Ashley</b> ?	(1) Addressing oneself, other people or things
40	tommy	20	104.11	1. Ventri was the last person to see <b>Tommy</b> alive. 2. ...if he didn't kill <b>Tommy</b> , who did?	(1) Addressing oneself, other people or things
41	know	218	263.95	1. We <b>know</b> she was upset with her parents. 2. I <b>know</b> how to use it...	(2) Informal spoken language use
42	here	159	251.95	1. So I probably shouldn't be standing around <b>here</b> ... 2. <b>Here</b> , I bought you a coffee...	(2) Informal spoken language use
43	yeah	162	210.80	1. <b>Yeah</b> . 2. <b>Yeah</b> , you do.	(2) Informal spoken language use
44	right	208	323.08	1. All <b>right</b> . 2. ...Paul always backed the van up <b>right</b> here. 3. He looked <b>right</b> into mine. I looked <b>right</b> into his eyes. 4. He's there <b>right</b> now. 5. Sharing information with the <b>right</b> person...	(2) Informal spoken language use
45	cop	25	186.30	1. Benny invited an undercover <b>cop</b> into the crew? 2. I wouldn't be sitting in a <b>cop</b> car... 3. Robbing a <b>cop</b> with no gun.	(2) Informal spoken language use
46	guys	31	184.76	1. <b>Guys</b> at the security desk know him as Jason... 2. I'm sure you recognize these <b>guys</b> .	(2) Informal spoken language use
47	tell	88	177.33	1. Go ahead, <b>tell</b> 'em. 2. Well, <b>tell</b> us.	(2) Informal spoken language use
48	let	78	158.70	1. <b>Let</b> me guess. 2. <b>Let</b> me ask you a question if the record. 3. <b>Let</b> me see. 4. <b>Let</b> 's see. 5. <b>Let</b> 's start with the stolen car.	(2) Informal spoken language use
49	huh	21	142.68	1. That's a lot of chocolate, <b>huh</b> ? 2. Man, he looks tired, <b>huh</b> ?	(2) Informal spoken

					language use
50	looks	52	142.53	1. <b>Looks</b> like he's putting the gun in your hand. 2. <b>Looks</b> like a burner. 3. Somebody <b>looks</b> like they're in a hurry.	(2) Informal spoken language use
51	like	177	112.27	1. Looks <b>like</b> a burner. 2. ...one of 'em looked <b>like</b> our vic. 3. How'd she afford a dress like that?	(2) Informal spoken language use
52	get	134	110.75	1. This cabdriver parked here to <b>get</b> a cup of coffee. 2. <b>Get</b> away from me. 3. I <b>get</b> there.	(2) Informal spoken language use
53	hell	31	100.49	1. What the <b>hell</b> are you doing, Frank? 2. Where the <b>hell</b> have you been? 3. What the <b>hell</b> is wrong with you?	(2) Informal spoken language use
54	murder	54	221.53	1. ...the champagne bottle become a <b>murder</b> weapon... 2. There was no signs of a struggle or <b>murder</b> . 3. Rennick's not getting away with <b>murder</b> . 4. You're under arrest for the <b>murder</b> of Jeremy...	(3) Crime scenes
55	killer	36	215.63	1. Our <b>killer</b> must've used Ashley's phone... 2. Is he our <b>killer</b> or our witness? 3. Nothing about these two murders suggests serial <b>killer</b> . 4. Bridge surveillance shows the <b>killer</b> was lying in wait... 5. Ellen may have known her <b>killer</b> .	(3) Crime Scenes
56	found	117	196.20	1. We <b>found</b> a bag of pills in his backpack. 2. I <b>found</b> a bloodstain on our victim's dress.	(3) Crime Scenes
57	kid	35	196.11	1. Except the <b>kid</b> has a clean record and his business was picking in. 2. A <b>kid</b> robs a jewelry store... 3. That <b>kid</b> , he pointed a gun and he shot at me...	(3) Crime Scenes
58	blood	56	175.14	1. ...I found a single <b>blood</b> drop, footprints... 2. Also found <b>blood</b> trace on a jewelry clasp. 3. I mean, but the <b>blood</b> pool is right here.	(3) Crime Scenes
59	in	596	137.63	1. I found this <b>in</b> his pocket. 2. Mary Portico is somewhere <b>in</b> San Francisco.	(3) Crime Scenes
60	out	227	133.31	1. ...Keith checked <b>out</b> a gun... 2. ...she made it <b>out</b> of New York...	(3) Crime Scenes
61	killed	44	129.92	1. We're not sure the rope <b>killed</b> her. 2. He was <b>killed</b> somewhere else, then	(3) Crime Scenes

				dumped here. 3. She was <b>killed</b> with a rock.	
62	kill	35	129.86	1. It was a trap designed to <b>kill</b> a firefighter. 2. Brooks used basic chemistry to <b>kill</b> Jimmy Clark. 3. Plenty of motive to <b>kill</b> . 4. Same gun used to <b>kill</b> the Jane Doe in the Hell's Kitchen.	(3) Crime Scenes
63	all	245	125.60	1. ... <b>all</b> transportation headed out of the city. 2. ... <b>all</b> with blood on their hands.	(3) Crime Scenes
64	shot	42	125.17	1. ...a little girl was <b>shot</b> and killed with your gun. 2. Jane Doe <b>shot</b> in Hell's kitchen 24 hours ago. 3. Man was <b>shot</b> in the back.	(3) Crime Scenes
65	no	242	120.98	1. <b>No</b> sign of Rachel. 2. ... <b>no</b> record of that name before 2000. 3. ... <b>no</b> background info, <b>no</b> permit...	(3) Crime Scenes
66	want	100	119.14	1. I <b>want</b> a name and address as soon as possible. 2. Her parents <b>want</b> her found, Oliver.	(3) Crime Scenes
67	guess	28	114.08	1. I <b>guess</b> this explains the drug trace in the van... 2. <b>Guess</b> whose name showed up on a list...	(3) Crime Scenes
68	name	64	108.73	1. The vic's <b>name</b> is Theodore Hart. 2. I'm gonna need that patient's <b>name</b> .	(3) Crime scenes
69	right	208	323.08	1. All <b>right</b> . 2. ...Paul always backed the van up <b>right</b> here. 3. He looked <b>right</b> into mine. I looked <b>right</b> into his eyes. 4. He's there <b>right</b> now. 5. Sharing information with the <b>right</b> person...	(4) Interviewing suspect, witness, or person
70	which	31	180.36	1. <b>Which</b> alley could he have dumped the gun in? 2. <b>Which</b> would mean the luxury sedan carjacked the van? 3. <b>Which</b> way?	(4) Interviewing suspect, witness, or person
71	why	106	156.07	1. <b>Why</b> 'd you book a room tonight at the St. Monarch hotel? 2. So, <b>why</b> didn't you call the cops? 3. <b>Why</b> hang somebody who's already dead? 4. <b>Why</b> robs someone and not take all the money? 5. <b>Why</b> would someone send those photos to her parents?	(4) Interviewing suspect, witness, or person