

Identifying Specialized Vocabulary in Thai Food Menus Using Computer-Based Approach

Piyada Low

Faculty of Management Sciences, Kasetsart University, Thailand

Email: fmspyl@ku.ac.th

Abstract—The primary aim of this study is to identify a specialized vocabulary in Thai food menus using a computer based approach. The study started from searching for Thai restaurant menus available online and downloading the menus from the webpages. The downloaded food menu items were changed into the plain text format (.txt) files and saved as the corpus for the study. The corpus analysis was performed with a lexical analysis software program, Lexical Tutor (Vocabulary Profile) which is an online tool. Corpus counting and word frequency lists were performed and presented in four vocabulary levels: high frequency words or the General Service List (GSL) with K1 (1-1000 word families) and K2 (1001-2000 word families) levels; the 570-word Academic Word List (AWL); and off-list words or specialized vocabulary. The results showed 55.27% of GWL (K1+K2), 0.89% of AWL, and 43.83% of off-list words which indicated the specialized vocabulary of the food menu corpus.

Index Terms—computer-based approach, corpus, specialized vocabulary, Thai food menus

I. INTRODUCTION

Vocabulary knowledge is important in language learning. [1] mentioned that vocabulary knowledge is a way that learners can improve second language proficiency. The numbers of words as well as which words needed and worth in very limited instructional time is important. Language teachers, therefore, would need research-based, empirically substantiated lists of the most important vocabulary words that language learners need to know.

The present study was conducted as a case study on how to utilize the Internet and a computer corpus program for systematic study of vocabulary to investigate the specialized vocabulary word list of Thai food menus rather than create a complete list of Thai food menus. Specialized or technical vocabulary would be of interest and the use of people working in a specialized field [2] as they are authentic examples [3], in this case, the specialized field is Thai cuisine.

Thai food has been promoted by the Thai government continuously. Final report of a survey on tourists' attitudes and satisfactions travelling Thailand in 2016 by Ministry of Tourism and Sports [4] revealed that Thai food was mentioned as one of the main five reasons for visitors to Thailand since 2012. The full report on foreign

visitors' attitudes and satisfaction travelling in Thailand in 2016 showed that Thai food has been rated as the second top reason to come to Thailand with 67.6% in 2015 and 66.8% in 2016. Thai restaurants as well as Thai cooks have to work hard to maintain the present situations of food tourism in Thailand. In order to perform tasks up to qualified standards, Thai cooks must have skills and knowledge in English as it is indicated as important knowledge for Thai cooks according to occupational standards and professional qualifications of Thailand Professional Qualification Institute [5]. Specialized food vocabulary word list in English, therefore, should be identified as the list will be beneficial and a valuable tool for Thai cooks and people related to Thai cuisine. As the study aims to identify the specialized vocabulary of Thai food menus, it tries to answer the following question:

What is the specialized vocabulary word list in Thai food menus?

Based on [6], vocabulary can be divided into four levels: high frequency words or the General Service List (GSL) which is commonly referred to as the 2,000 most frequent words in English; the 570-word Academic Word List (AWL) which are words mostly found across a wide range of academic texts; specialized or technical vocabulary which are words related to specific areas such as engineering, medicine, linguistics, etc.; and the low frequency words which are the remaining words of English. These word frequency lists would help language learners to determine the most useful words to study [7]. Implications for teaching English for Specific Purposes (ESP) and specialized vocabulary in other fields are recommended based on the study approach.

II. REVIEWED LITERATURE

A. Specialized Vocabulary

[8] explained that specialized vocabulary as words that were "recognizably specific to a particular topic, field or discipline". If specialized vocabulary were found as much as 30% of a text, it could be difficult for students to acquire these words incidentally when they first met them and it would be slow and frustrating to use dictionary constantly [2]. There were researchers gave the importance on the value of ESP classes which concentrating on the vocabulary that the students need for their disciplines, and corpus-informed lists should be

established from the texts that students will need to read and the genres that they will need to write in for their various disciplines [2], [3], [8], [9].

B. Related Research Studies

To teach and learn a language, the amount of words and word selection for teaching in limited instructional time is important. Research-based, empirically substantiated lists of the most important vocabulary words would be needed for language learners. Corpus linguistics, therefore, allows teachers to produce vocabulary lists that better match learners' actual English needs [1]. Research studies on vocabulary knowledge and corpus have been conducted in various fields of study such as anatomy, linguistics, television shows, newspaper reviews, culinary course, marketing, wine tasting course, and English for Specific Purposes (ESP) course. Table I shows research studies on vocabulary knowledge and corpus.

TABLE I. RESEARCH STUDIES ON VOCABULARY KNOWLEDGE AND CORPUS

Contents	Results
[2] Anatomy text and linguistics text	Anatomy text mostly has technical vocabulary of which 64.4% is restricted to the field of anatomy. Applied linguistics has less technical vocabulary of which 88.4% is words generally familiar to people. The most obvious technical words are those with Greek or Latin based forms which occur only in the specialized area.
[10] Television Word List	Ten TV shows, five comedies and five dramas, were used to create the corpus of the study. Few AWL (29.46%) appeared in the corpus while 70.54% occurred in off-list words. Informal words or spoken words were mostly found in the corpus. 'Okay' and 'guy' are the most frequently and consistently used words in the television word list.
[11] Newspaper Restaurant Reviews	Corpus and data were randomly selected from restaurant reviews from five leading newspapers in the US in 2010. Results show that most of the words belong to the GSL words, and 23.21% of the words were specialized vocabulary and considered as authentic materials for ESP learning.
[12] Culinary Course	The compilation of 11 PowerPoint slide presentations in classes were used as the corpus. Results show 113 specialized vocabulary. Words from the corpus were categorized according to the rating scale. Ingredients/food/drink is mainly rated. However, the researchers found out that the amount of vocabulary was relatively too small for vocabulary development.
[13] Marketing	The compiled corpus was taken from four written texts: books, journals, websites, and newspapers. There were 3,258 high frequency content words from the whole corpus. The corpus consisted of 1,476 running words found in GSL, 837 words found in AWL and 468 high frequency technical words.
[3] Wine Tasting Course	Two in-house corpora were compiled and used as supplementary materials for the students' practice and homework. The analysis showed that more specialized corpora tend to have a higher frequency of content words as reflected in the high lexical density (greater than 66%) and the strong presence of specialized vocabulary (more than 24%).
[14] English for Specific Purposes (ESP) Course	Two food groups were chosen as needed for vocabulary development: 1. herbs, spices, and food seasonings (98.1%) 2. fish and shell fish (96.3%). Students indicated that fish and shell fish was the

	great problem for them (48.1%). More than half of the students (63%) admitted that task assignment responded to their needs of vocabulary knowledge development.
--	--

In conclusion, vocabulary knowledge is important for non-native English learners, especially specialized vocabulary and it should not be a barrier for professional practice [15] because vocabulary learning is teachable [16]. Limited resource from textbooks will be too easy for students and will not enhance vocabulary learning whereas multiple exposures to target words help accelerate vocabulary growth [12] and specialized vocabulary should be implemented for ESP course [3]. Authentic materials would be useful for language learning, especially vocabulary [2], [3].

III. METHODOLOGY

The methodology of this research can be divided into three parts: corpus collection, corpus processing, and corpus analytical tools.

A. Corpus Collection

Thai Food menu items which are authentic materials were used in the research study. Authentic materials enhance language acquisition and cultural awareness [17]. The food menu items were downloaded from six Thai restaurants webpages and saved in Microsoft Word Office 2013. The main principle is that these Thai restaurants must be in English speaking countries [18] so they can represent the Thai food menus that native English speakers would understand. The six English native speaking countries are Australia, Canada, Ireland, New Zealand, UK, and US. The selected six Thai restaurants are Tawandang, Australia; Pai Restaurant, Canada; Saba Restaurant, Ireland; Suk Jai Thai Restaurant, New Zealand; Chaophraya, UK; and Nahm Thai Cuisine, US. The size of the food menu corpus contains 7403 tokens or running words.

B. Corpus Processing

The downloaded Thai food menu items were standardized before being used as the corpus. Numbers, Thai words, English transliteration of Thai words, proper names, names of the restaurants and other components that could not be processed by the computer programs and were not parts of lexical analysis were removed. The food menu items, then, were converted to the plain text format (.txt) files for the computer program to run over.

C. Corpus Analytical Tools

Lexical Tutor (Vocabulary Profile) is a lexical analysis software program used for corpus analysis in this study. The program is available free on the Internet. It is an online tool used to analyze and categorize the corpus into General Service List (GSL), Academic Word List (AWL), and off-list words (<https://www.lexutor.ca/vp/>). The results show word lists into four levels. Level 1 is K1 words, list of 1-1,000 word families including both function and content words; Level 2 is K2 words, list of 1,001-2,000 word families. These two lists are considered as general words used generally in both spoken and written

texts. Level 3 is for academic words consisting of 570 word families outside the GSL and used in academic texts. Level 4 is off-list words or words that do not occur in K1, K2, and AWL lists. The off-list words refer to specialized vocabulary related to specific fields such as linguistics, anatomy, television, newspaper reviews, culinary course, marketing, etc. These three base lists served as basic corpus processing units in this corpus study.

IV. RESULTS

The vocabulary profile of the corpus presented in Table II was based on the Lexical Tutor analysis. The GSL in K1 words appeared 35.47% and K2 words appeared 19.80% of the text. This means that the GSL (K1+K2) appeared most at 55.27% in the food menu corpus. These two word lists are commonly referred to as the 2,000 most frequent words in English. Only a small coverage (0.89%) of the words appeared on the AWL or words mostly found across a wide range of academic texts.

However, almost half of the vocabulary of the corpus (43.83%) consisted of off-list words or specialized vocabulary. This high percentage indicates that specialized vocabulary plays an important role in the food menu corpus and should be focused in this ESP course.

TABLE II. VOCABULARY PROFILE OF THE CORPUS IN FOUR LEVELS

	Tokens	Percentage
K-1 Words (1-1000)	2626	35.47%
K-2 Words (1001-2000)	1466	19.80%
AWL Words	66	0.89%
Off-List Words:	3245	43.83%
Total (unrounded)	7403	≈100.00

Based on the Lexical Tutor analysis, small coverage of the AWL words or 66 tokens (0.89%) appeared in the corpus. Though the AWL is small in coverage, the word list indicates the characteristics and nature of Thai cuisine. The first three words in the AWL are 'style, traditional, and selected'. These three words emphasizes that these Thai restaurants give importance on Thai style, tradition, and culture together with selected ingredients in cooking Thai food in order to enhance the restaurants' standard. Ranks of the 66 AWL words were presented in Table III.

TABLE III. RANK OF THE AWL WORDS IN THE CORPUS

Rank	AWL Words	Rank	AWL Words
1	style	6	create
2	traditional	6	created
3	selected	6	medium
4	classic	6	selection
5	enhanced	6	uniquely
5	items	7	features
5	minimum	7	substituted
		7	ultimate

From 3,245 tokens of off-list words, the 20 most frequent specialized vocabulary in the food menu corpus analyzed by Lexical Tutor was listed in Table IV. The vocabulary are mostly of cooking ingredients, tastes, and food textures.

Three specialized vocabulary occurred more than 100 times in the list are 'chili', 'curry', and 'onion' which are regular cooking ingredients of Thai cuisine.

TABLE IV. THE 20 MOST FREQUENT SPECIALIZED VOCABULARY

Rank	Word Type	Frequency	Rank	Word Type	Frequency
1	chili	145	9	pork	62
2	curry	129	12	coconut	61
3	onion	104	13	crispy	59
4	basil	80	14	peanuts	57
4	noodles	80	15	prawn	56
6	spicy	72	16	beef	52
7	vegetables	70	16	salad	52
8	shrimp	69	18	garlic	48
9	lime	62	19	peppers	45
9	mushrooms	62	20	tamarind	36

V. DISCUSSION AND CONCLUSION

The study aimed to identify the specialized vocabulary in Thai food menus. The food menu corpus contained 7403 running words. Most of the words appeared in GSL (K1+K2) at 55.27% while AWL appeared only 0.89%. Off-list words was relatively high at 43.83% in the food menu corpus. The result of the high frequency of off-list words is in line with the findings on specialized vocabulary of [2] in anatomy text and linguistics text, [10] in TV shows, [3] in wine tasting course, and [10] in newspapers restaurant reviews.

According to [2], when specialized vocabulary occurred as much as 30% of the text, it would be difficult to the language learners to acquire the words. Therefore, the off-list words could cause difficulty in learning process of nonnative speakers of English. Specialized vocabulary word list from the study can help learners study more effectively [19].

'Chili', 'curry', and 'onion' are most frequent specialized vocabulary occurring more than 100 times in the food menu corpus. Chili, onion, and basil are the top three vegetables in the specialized vocabulary. Most of the words in the specialized vocabulary list are ingredients in cooking or food names. These findings can be explained as follows:

1. Chili is mostly found in Thai dishes of different kinds: snacks, soups, salads, fried dishes, steamed dishes, etc.
2. Curry is in daily meals of Thai cuisine. Curry with steamed rice can be for breakfast, lunch, and dinner.
3. Onion is widely used in a variety of Thai dishes especially Thai spicy salads.

Moreover, eleven word types of the 20 most frequent specialized vocabulary are fruits and vegetables. This

finding is in line with [20] who pointed out that vegetables are important in Thai cuisine though they are often mixed with meat, poultry or fish and eaten as salad. The fresh ingredients and high nutritional value are mentioned as uniqueness of Thai food and enhance the reputation of Thai cuisine as one of the best in the world [21].

However, two words indicating food flavor and texture were found: 'spicy' and 'crispy'. The results were in line with the fact that spices and herbs are fundamental elements of Thai cuisine. Spicy flavor is common and like a signature of Thai food although not all Thai food dishes are spicy. For [20], a fresh sweet-sour taste is typically Thai. It is not a surprise that the words related to the tastes of spicy and sour are on the list: 'chili', 'spicy', 'lime', 'sour', and 'tamarind'. It is definite that these specialized words represent Thai food culture and the tastes of Thai food.

The results from the specialized corpus analysis can be used as inputs for English for Specific Purposes (ESP) course and future research studies in related field. The food menu corpus will contribute to the revised contents of English for Food and Beverage Service course which is a compulsory course for students majoring in Hotel and Tourism Management at the Faculty of Management Sciences, Kasetsart University Si Racha Campus, Thailand.

ACKNOWLEDGMENT

This paper was financially supported by the Faculty of Management Sciences, Kasetsart University.

REFERENCES

- [1] A. Youngblood and K. Folse, "Survey of corpus-based vocabulary lists for TESOL classes," *MexTESOL Journal*, vol. 41, no. 1, pp. 1-15, 2017.
- [2] T. M. Chung and P. Nation, "Technical vocabulary in specialised texts," *Reading in a Foreign Language*, vol. 15, no. 2, pp. 103-116, 2003.
- [3] H. Hou, "Teaching specialized vocabulary by integrating a corpus-based approach: Implications for ESP course design at the university level," *English Language Teaching*, vol. 7, no. 5, pp. 26-37, 2014.
- [4] Final report of a survey on tourists' attitudes and satisfactions travelling Thailand in 2016, Ministry of Tourism and Sports. (2017). [Online]. Available: <http://secretary.mots.go.th>
- [5] Occupational standards and professional qualifications of Thailand, Professional Qualification Institute. (2017). [Online]. Available: <http://tpqi-net.tpqi.go.th>
- [6] I. S. P. Nation, *Learning Vocabulary in another Language*, Cambridge: Cambridge University Press, 2001.
- [7] J. Spiri, "Online study of frequency list vocabulary with the WordChamp website," *Reflections on English Language Teaching*, vol. 7, no. 1, pp. 21-36, 2008.
- [8] P. Nation. (2001). *Learning vocabulary in another language*. Cambridge: CUP. [Online]. Available: <http://dx.doi.org/10.1017/CBO9781139524759>
- [9] C. Sutarsyah, P. Nation, and G. Kennedy, "How useful is EAP vocabulary for ESP? A corpus based study," *RELC Journal*, vol. 25, no. 2, pp. 34-50, 1994.
- [10] K. MacFadden, K. Barrett, and M. Horst, "What's in a television word list? A corpus-informed investigation," *COPAL*, vol. 2, pp. 78-94, 2009.
- [11] H. Hou, "A corpus approach to discourse analysis of newspaper restaurant reviews: A preliminary analysis," *Studies in Literature and Language*, vol. 5, no. 3, pp. 95-100, 2012.
- [12] N. R. M. Nordin, S. H. Stapa, and S. Darus, "Developing a specialized vocabulary word list in a composition culinary course through lecture notes," *Advances in Language and Literary Studies*, vol. 4, no. 1, pp. 78-87, 2013.
- [13] S. Kakaew and A. Tongpoon-Patanasorn, "Wordlists of marketing: a corpus-based study," presented at 9th International Conference on Humanities and Social Sciences, Khon Kaen, 2013, pp. 1668-1677.
- [14] P. Low, "Students' needs and vocabulary development in english for specific purposes (ESP) course," *Journal of English Language Teaching and Linguistics*, vol. 3, no. 2, pp. 151-161, 2018.
- [15] M. Diaz-Gilbert, "Vocabulary knowledge of pharmacy students whose first or best language is not English," *American Journal of Pharmaceutical Education*, vol. 68, no. 4, article 91, 2004.
- [16] V. Jurkovic, "Vocabulary learning strategies in an ESP context," *Scripta Manent*, vol. 2, no. 1, pp. 23-32, 2006.
- [17] J. Beresova, "Authentic materials- enhancing language acquisition and cultural awareness," *Procedia-Social and Behavioral Sciences*, vol. 192, pp. 195-204, 2015.
- [18] List of majority native English speaking countries. The University of Sheffield, UK. [Online]. Available: <https://www.sheffield.ac.uk/international/english-speaking-countries>
- [19] J. Liu and L. Han, "A corpus-based environmental academic word list building and its validity test," *English for Specific Purposes*, vol. 39, pp. 1-11, 2015.
- [20] G. Rosalin and T. I. Soetanto, "Analysis of the customers' preferences in determining a suitable menu: A case study of river side oriental cuisine restaurant," *Journal Manajemen Perhotelan*, vol. 2, no. 2, pp. 57-67, 2006.
- [21] E. V. Murray, "Thailand-The kitchen of the world: Origin and growth of the Thai Food Industry & Lessons for India," *Cab Calling*, April-June 16-26, 2007.



Piyada Low was born in Bangkok, Thailand. She graduated with BA in English Language and Literature from Thammasat University, Thailand and MA in English for Communication from Burapha University, Thailand. She also holds diplomas in hotel management from American Hotel and Motel Association and Hotel Institute Montreux, Switzerland.

She started her career as an ESL teacher for UNHCR at a refugee camp in Phanat Nikhom, Chon Buri, Thailand. She is currently an assistant professor at the Faculty of Management Sciences, Kasetsart University Si Racha Campus, Thailand.