Corpus-Based Approach to Generate a Word List for Food Service

Thanawan Rungrueang, Pitchapha Boonprasert, Sitanan Poempongsajaroen,

Piyapong Laosrirattanachai*

*Corresponding author's email: piyapong.l@ku.th

Kasetsart University, Thailand

Abstract

Learners preparing to work in the food service industry should be equipped with food service terminology to enhance their communicative potential. The purpose of this study was to create a Food Service Word List (FSWL) to assist L2 learners of English in English for Specific Purposes classrooms. The Food Service Corpus (FSC) with a size of 1,871,271 running words was created by collecting data from four websites relevant to the food service industry, consisting of www.bbc.co.uk/food/cuisines, www.tasteatlas.com, www.escape.com.au and www.edition.cnn.com. The procedure of generating the FSWL comprised three stages: keyword analysis was used to extract the first 814 words with the highest LL values; lexical profiling was used to select words allocated outside the referent word lists; and three experts in the food service industry were asked to give their viewpoints based on their experiences. The results produced the FSWL containing 261 words with a coverage of 8.64% of the FSC. The curated word list can enhance the food service vocabulary range of learners and prepare them for more effective communication in a food service career.

Keywords: food service, corpus linguistics, word list, Food Service Word List, specialised vocabulary

Introduction

English is one of the most extensively used languages in the world. Nowadays, English is the most essential communicational tool and continually increases its impact as a tool of communication for people around the world both in local and global contexts. Communication in English requires many elements including listening, speaking, reading, writing, grammar, and vocabulary. Among these skills, vocabulary is the most fundamental element of language communication (Laufer & Nation, 1995; Lestari & Hardiyanti, 2020). One of the most famous ideas emphasising the importance of vocabulary was that even if we lack grammar knowledge, we are still capable of communicating, but if we lack vocabulary knowledge, we hardly communicate comprehensibly (Nosratinia et al., 2013; Wilkins, 1972). For this reason, vocabulary is considered as the heart of a language and a key to communication. Having a knowledge of vocabulary makes all communication skills easier to perform (Attachoo &

Chaturongkul, 2015; Paakki, 2013). Constructing a word list is one way to help people increase their vocabulary size using the corpus linguistic method, especially for specialised words in a specific field.

There are many technical words used in daily life depending on the topic. The food industry is one specific area that definitely has a considerable impact on people worldwide because food is one of four requisites necessary for living. In fact, everyone needs to consume at least one-to-three meals per day for healthy living (Soonthornsaratool, 2018). Food also is important in various ways, such as culture, tradition, and economics. Food represents the cultural identity of each nation. Traditional cuisine is handed down from generation to generation. People from different cultures eat different foods. There are many possible factors that affect food attributes, for example belief, way of life, art and culture, tradition, and local wisdom. As a result, different ingredients, cooking processes, and ways to present dishes lead to food diversity. People are becoming more interested in food from different cultures and nowadays tourists are embarking on culinary journeys to taste the dishes from different cultures globally. Karim and Chi (2010) claimed that one of the interesting trends of tourism around the world is 'gastronomy tourism.' Famous countries that attract many tourists with their authentic cuisines include France, Italy, and Thailand. The food service industry plays an important role in the economy as well as it generates a lot of income in the country. Food Intelligence Center (2020) conducted a survey concerning income and Thailand's Food Industry. The survey showed that in 2019, food industry revenue was USD 29,479 million, approximately 5.5% of Thailand's GDP and 20.6% of Thai industrial GDP. The thriving food service industry increases the number of new restaurants, such as cafés, drive-thru restaurants, food trucks, and online restaurants with a delivery service. This phenomenon leads to additional employment in this line of work, especially for restaurant workers, who are responsible for providing service in the food service industry, and other positions related to the industry.

It is necessary for the people working in such an industry to have an adequate knowledge of the English language, especially of the vocabulary used in the food service industry to respond to the demands of service users. Preparing learners who intend to enter the food service industry to be able to communicate effectively is also an important task of teachers in the field. Effective and smooth communication requires understanding and using technical words in the food service industry for describing cooking recipes, explaining dishes, and recommending menu items. Therefore, the present study aimed to systematically create a Food Service Word List (FSWL) to facilitate learners intending to enter the food service industry. The corpus method was utilised, and three main criteria (keyword analysis, lexical profiling, and expert viewpoints) were applied to generate the FSWL with the expectation that it would enhance the vocabulary of learners.

Theoretical Background

Corpus Linguistics

A corpus is a group of texts collected and stored in an electronic database used for linguistic analysis and constructed for a specific purpose. A corpus is a data-driven tool for learning a language (Aroonmanakun, 2011) where it is easy to access the data using a computer and to analyse the data using specialised computer software programmes. O'Keeffe and McCarthy (2010) claimed that corpus linguistics has been significant in several fields, for example learning and teaching language and vocabulary, discourse analysis, pragmatics, forensic linguistics, speech technology, and sociolinguistics. The evolution of corpus linguistics produces new theories of language, while a corpus can play an important role in English language teaching. Corpus linguistics is described as the study of language that exists naturally which enables learners to use the language that is applied in actual life (Crystal, 1991). In the vocabulary research field, corpus linguistics has played a major role as a reliable and systematic method facilitating research. For example, lexicographers use a corpus to discover frequently used words and then apply this information to categorise important headwords for teaching and learning and include them in learners' dictionaries (Garcia, 2014). Teachers apply the corpus linguistics method to indicate the most frequent words and phrases that occur in English to be included in the lesson content (Bennett, 2010). The study of specialised word lists is an important part of corpus linguistics in the field of English for Specific Purposes (ESP) (Kruawong & Phoocharoensil, 2020). Therefore, corpus linguistics was used in the present study to generate the FSWL in the present study.

Word List Background and Category of Word List

Traditionally, a word list is a list of vocabulary that is usually classified by its frequency order or in alphabetical order (O'Keeffe et al., 2007). Later, a word list was mainly constructed to gather vocabulary frequently used in a specific genre. A word list is valuable in several linguistic studies, such as language teaching (Gardner & Davies, 2014). Nation (2013) classified vocabulary into four categories for learning the English language consisting of high-frequency words, academic words, technical words, and low-frequency words, which can be further explained as follows.

The high-frequency words are words that we frequently encounter in daily life. The most famous word list representing high-frequency words is the General Service List (GSL). The GSL is a list of approximately 2,000 high-frequency word families created by Michael West in 1953. Later in 2013, Browne modified the list and proposed the New General Service List (NGSL), claiming that the first 1,000 highest frequency words in the NGSL covered up to 90% of most general English texts.

The second category of word list is the academic word list. The most recognisable academic word list is that of Coxhead (2000) called the Academic Word List (AWL). Coxhead gathered language

data from different genres of academic texts including linguistics, science and economics. There are approximately 570 word families in the list. The AWL was claimed to cover around 10% of academic texts, usually for academic purposes (Coxhead, 2000).

The third category of word list is a technical word list. Technical words refer to vocabulary used in a particular field, whose uses have specific meanings. Technical words cover about 5% of specialised texts. Recently, this category of word list has gained enormous attention from teachers since they aim to create technical word lists to serve their students' needs.

The last category of word list is a low-frequency word list that contains vocabulary that occurs infrequently and is not included in the GSL, AWL, or technical word lists. Nation (2013) further explained that low-frequency words could be technical words when appearing or being used in other fields.

Technical Word Lists

A technical word list can be variously called a specialised word list, field-specific academic vocabulary list, discipline-specific academic word list, and discipline-based lexical repertoires. A technical word list is a list of vocabulary that is used in particular fields and for particular purposes (Liu & Han, 2015). Nation (2016) explained that there are two types of specialised words: those normally known by the general public and those known only by specialists. According to Laufer (1989), to read a text comprehensively, one should know vocabulary covering about 95% of the text. Generally, the GSL covers about 80% of texts (Nation & Waring, 1997), and the AWL covers about 10% of texts (Coxhead, 2000). Therefore, 5% of the text coverage should be technical words. However, some studies have shown that the text coverage of different technical word lists may vary according to the field (Chung & Nation, 2003; Coxhead et al., 2020; Hyland & Tse, 2007). There are various possible factors affecting the proportions of components in the word list coverage, such as the sources used to compile the corpus and the methods used to construct the word list. For example, the lexical profiling method (as explained in the next section) could eliminate words identified as being high-frequency. As a result, the size of the technical word list, especially a specialised word list comprising words commonly known by people, could be narrowed down. Thus, the text coverage could be reduced according to the narrowed-down word list. There are many benefits from creating a technical word list; for example, it facilitates learners to expand the size of their vocabulary in a specific or technical area, can be used as a teaching and learning material and applied in English for Specific Purposes courses, and can be diversified and adapted to be suitable for different learning strategies (Nation, 2006). Technical word lists have been of interest for many years. Many scholars have tried to create and develop technical word lists. See Table 1.

Table 1

Some Examples of Technical Word Lists

Category of technical word list	Creator
Agriculture	Martinez et al. (2009), Muñoz (2015)
Business	Konstantakis (2007)
Chemistry	Valipouri and Nassaji (2013)
Culinary	Nordin et al. (2013)
Engineering	Ward (2009), Watson Todd (2017)
Environment	Liu and Han (2015)
Hospitality (Tourism, Hotel, and Airline)	Laosrirattanachai and Ruangjaroon (2021)
Nursing	Yang (2015)

In the present study, we aimed to generate a Food Service Word List as a list of words specifically used in the food service discipline. In Table 1, two of the word lists are related to food service: the Hotel Business Word List (HBWL) (Laosrirattanachai & Ruangjaroon, 2021) and the Composition Culinary Course Word List (CCCWL) (Nordin et al., 2013). The HBWL, constructed from the corpus compiled from official hotel websites and hotel business news, contains 274 words. It was created to help learners in the field of hotel business to enhance their vocabulary ability. The CCCWL, created from a compilation of all PowerPoint slides from one writing course, contains 113 words. However, the benefit of using each word list in a food service course needs to be further investigated. There are many vocabularies related to the food service industry. If learners do not know or do not study these specific words in food service, there may be a lack of comprehension and communication barriers.

Constructing a Word List

To generate a word list, there are five main criteria widely used by many scholars consisting of frequency, range, lexical profiling, keyword analysis and expert viewpoints (Coxhead, 2000; Lei & Liu, 2016; Laosrirattanachai & Ruangjaroon, 2020; Muñoz, 2015; Tangpijaikul, 2014; Watson Todd, 2017; West, 1953; Yang, 2015).

Frequency is the most classic and extensively used criterion for language learning (Lindquist, 2009). However, frequency is not sufficient to generate a reliable word list because it might be biased by longer texts (Coxhead, 2000).

Thus, range became a crucial tool to reduce the size and remove bias in the word list. Range is used to investigate the number of sources in which a word appears. According to Nation (2016), a word may occur very often in one source, but not appear in other sources. Such a word is labelled as having a narrow range. A word with a high range value is considered for the list.

Lexical profiling is used to remove words that are not relevant in the field by eliminating words that appear in a referent word list such as the GSL (West, 1953), AWL (Coxhead, 2000), and other word lists. Words that do not appear in the referent word lists are normally considered to have passed this criterion and are contenders for inclusion in the list. Coxhead (2000) used this criterion to remove words that occurred in the GSL to construct the AWL.

Keyword analysis is one criterion that is widely used in corpus linguistics and also to generate a word list by considering unusual frequency words that occur in a target corpus compared with the referent corpus (Gabrielatos & Marchi, 2012) based on log-likelihood value (LL). Keyword analysis can be used in diverse ways to create a word list. For example, Watson Todd (2017) chose to consider the relative frequency from the keyword analysis instead of considering the absolute frequency to create an engineering word list. Tangpijaikul (2014) substituted the keyword analysis for both frequency and range to create a business vocabulary list. Laosrirattanachai and Ruangjaroon (2020; 2021) applied keyword analysis as a tool to recall words that appeared in the GSL and AWL but have specific meanings in hospitality businesses to create the tourism, hotel, and airline business word lists.

Expert viewpoints are very useful in constructing a word list. Opinion and feedback from experts who have experience in specific fields can help to consider which words are appropriate to include in word lists (Chung & Nation, 2004; Martinez et al., 2009; Schmitt, 2010). Chung and Nation (2004) suggested using four rating scales to collect experts' viewpoints. Each scale was described as follows. The first scale refers to words with a meaning irrelevant to the food service field and such words scored as 1. The second scale, scored as 2, refers to words with a meaning of little relevance to the food service field. The third scale, scored as 3, refers to words with a meaning very relevant to the food service field. The fourth scale, scored as 4, refers to words with a meaning solely related to the food service area and not included in other fields.

There is no evidence confirming which criterion is the best to create a word list because each has different advantages and disadvantages. Hyland and Tse (2007) affirmed that each field has different restrictions for generating a word list. In the present study, three main criteria (keyword analysis, lexical profiling, and expert viewpoints) were used to generate the FSWL.

Research Questions

The purpose of this study was to create a Food Service Word List (FSWL) to assist L2 learners of English in English for Specific Purposes classrooms. The research questions of the study were:
1. What words should be included in the FSWL?
2. What is the proportion of Food Service Corpus coverage in the FSWL?

Research Instrument and Methods

Corpus Compilation

To generate the Food Service Word List (FSWL), we first developed the Food Service Corpus (FSC). The FSC was compiled by gathering language data related to the food service industry from four websites: www.bbc.co.uk/food/cuisines, presenting food from 22 nationalities around the world; www.tasteatlas.com, ranking the top-100 most popular dishes in the world, the top 100 most popular foods in the world, and the top 100 best related foods in the world; www.escape.com.au, providing information on the "99 things to eat that will blow your mind"; and www.edition.cnn.com, rating the world's 50 best foods. The gathered texts provided information on a brief history of the food, its ingredients and recipes, and how it should be served. Data were collected from July to August 2021. In total, the FSC contained 1,871,271 running words.

Software Tools and Research Instrument

In the present study, 2 software tools were used to generate the data: Key-BNC (Graham, 2021) and AntWordProfiler (Anthony, 2021), in conjunction with a questionnaire. Both programmes are available as freeware for corpus linguistics purposes. The Key-BNC programme was used to undertake keyword calculations and analysis based on the log-likelihood value (LL) of the words in the FSC compared with the British National Corpus (BNC). The Key-BNC programme categorises the words in order from the highest to the lowest LL value. The AntWordProfiler programme was used to allocate words into different referent word lists, based on profiles provided in the programme. Words allocated in the profiles were finally removed from the FSWL. Questionnaires containing a list of possible words were distributed to the experts. The questionnaires were designed based on a 4-rating scale applied from Chung and Nation (2004).

Data Analysis

According to Hyland and Tse (2007), the construction of different word lists requires different procedures. Furthermore, Laosrirattanachai and Laosrirattanachai (2021) claimed that different corpora required different criteria in constructing a word list. The FSC was compiled from language data concerning dishes around the world. Each dish somehow is unique and is different from others in some ways. It might be too difficult to rely solely on absolute frequency and range, since some words related to food service may have a mid-frequency rate or not appear in as many sources as expected. Furthermore, if a learner needs to study English for a specific purpose, such as food service, it is more beneficial to learn a list of specialised words that occurs more frequently in the food service context than to learn a list that occurs in a general English context (Gries, 2015). Accordingly, some word list studies adopt this concept and apply keyword analysis using relative frequency instead of absolute

frequency (Tangpijaikul, 2014; Watson Todd, 2017). Therefore, in the present study, we applied the idea of Tangpijaikul (2014) by using keyword analysis instead of frequency and range. To systematically generate the FSWL, keyword analysis, lexical profiling, and expert viewpoints, respectively, were used.

Figure 1

Construction Procedure of Food Service Word List
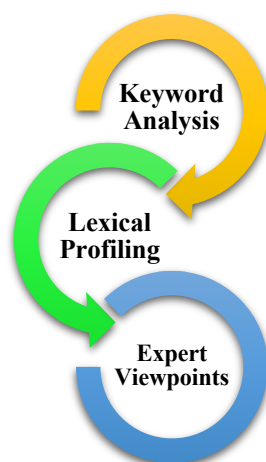
Keyword analysis was used as the first criterion to generate the FSWL. The cut-off point in the present study was adapted from Watson Todd (2017). In Watson Todd's study, the top-500 words with the highest LL values from the corpus with a size of 1,150,000 tokens were considered to be the cut-off point and this was applied in the present study using the following equations:

$$\frac{500}{1,150,000} = \frac{X}{FSC\ size}$$

$$X = \frac{500\ x\ FSC\ size}{1,150,000}$$

$$X = 813.59$$

The cut-off point for the present study was 813.59 that was rounded up to 814. Thus, the first 814 words with the highest LL values were considered as passing the keyword analysis criterion.

Lexical profiling was used as the second criterion after the process of keyword analysis in the present study. The 814 words from the first process were allocated into the referent word lists using AntWordProfiler (Anthony, 2021). Words allocated in six referent word lists comprising the first 1,000 GSL, second 1,000 GSL, AWL, Function Word List (FWL), Abbreviation List (AL), and the Proper Name List (PNL) were removed. The FWL, AL, and PNL can be obtained from https://www.wgtn.ac.nz/lals/resources (Nation, 2018). The remaining words were considered as possibly relevant to the food service industry and were considered by three experts as the last criterion used in the present study. To assure that the completed Food Service Word List was useful, reliable, relevant, and appropriate for learners in the food service field, the FSWL was verified by consulting

experts in the food service field. In the present study, three experts, who each had worked for more than four years at the assistant manager level in restaurants and used English to communicate and provide service to foreign customers, were asked to provide their viewpoints on our proposed FSWL. The word-list check lists with the four rating scales applying the criteria from Chung and Nation (2004) were distributed to the three experts. An explanation of the four rating scales was also provided in the questionnaire. Words scored three or four by at least two experts passed this criterion and were included in the final FSWL.

Results

After compiling the FSC consisting of 1,871,271 tokens, three criteria (keyword analysis, lexical profiling, and expert viewpoints) were used to refine this to the FSWL. The results for each criterion are presented below.

Keyword Analysis

Keyword analysis was used as the first step to generate the Food Service Word List using the Key-BNC programme. According to the study ratio by Watson Todd (2017), the first 814 words with the highest LL value met the cut-off point criterion. Some sample words passing the keyword analysis criterion with their LL values are shown in Table 2.

Table 2

Examples of Keyword Analysis based on Log-Likelihood Value (LL)

Word Type	Freq.	Freq. BNC	LL	Word Type	Freq.	Freq. BNC	LL
add	14,192	8,393	83,914.63	cook	6,791	3,887	40,407.40
tbsp	10,780	569	81,647.81	garlic	5,489	802	39,090.75
chopped	9,704	1,049	70,707.53	finely	5,376	726	38,532.91
salt	10,300	3,105	67,910.80	sugar	6,387	3,953	37,428.66
minutes	13,651	17,771	66,725.18	oz	5,455	1,217	37,296.80
until	16,547	40,058	65,309.23	oven	5,366	1,354	36,176.32
heat	10,721	6,040	63,983.74	serve	6,524	5,300	36,064.75
tsp	8,115	374	61,793.16	bowl	5,626	2,418	35,210.80
oil	11,323	10,776	60,258.80	flour	5,079	1,062	34,969.13
pan	8,686	2,347	58,075.81	stir	5,094	1,190	34,650.40
pepper	7,241	979	51,895.64	chicken	4,929	2,112	30,863.81
ingredients	6,820	1,243	47,612.06	freshly	4,248	718	29,868.06
sauce	6,562	1,432	44,973.74	mixture	5,067	3,216	29,543.72
butter	6,838	2,165	44,798.45	olive	3,948	991	26,632.97

According to Table 2, most of the illustrated words were relevant to the food service field, such as *chopped, salt, heat, ingredients, serve, stir, oven, pan, butter*, and *olive*. However, some words not directly relevant to the food service area also appeared in the list, such as *until, finely, freshly, mixture*, and *minutes*. The 814 words that passed the keyword analysis criterion were considered according to the next criterion of lexical profiling. Some sample words that passed the keyword analysis criterion with their contexts extracted from the FSC are shown below.

Thread 1: Stir

*The pasta must also be taken off the heat before the eggs are **stirred** in.*

Thread 2: Flour

*The simple dough is prepared with **flour**, yeast, salt, and lukewarm water.*
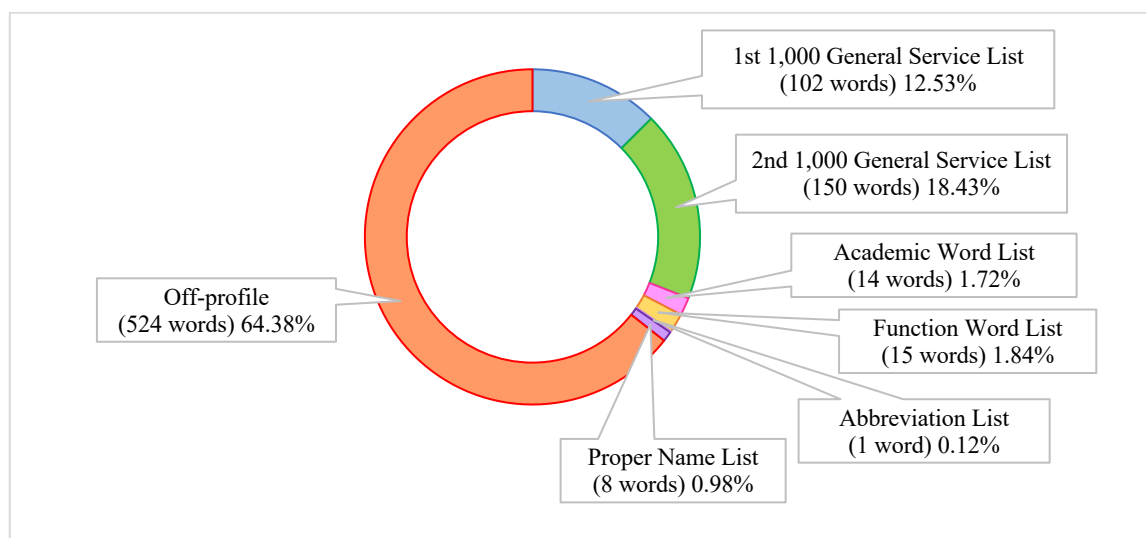
Thread 3: Heat

*Put the curry paste and chickpeas in a frying pan and cook over a low **heat** for 1–2 minutes, stirring constantly.*

Lexical Profiling

After generating the list of words arranged by their LL values from the highest to the lowest, the top 814 words were considered to have passed the first criterion and were then analysed by using the AntWordProfiler programme to remove the words that occurred in the GSL, AWL, FWL, AL, and PNL. The results are shown in Figure 2

Figure 2

Proportion of Words Occurring in Each Referent Word List

1st 1,000 General Service List
(102 words) 12.53%

2nd 1,000 General Service List
(150 words) 18.43%

Academic Word List
(14 words) 1.72%

Function Word List
(15 words) 1.84%

Abbreviation List
(1 word) 0.12%

Off-profile
(524 words) 64.38%

Proper Name List
(8 words) 0.98%

From Figure 2, the AntWordProfiler programme removed from the list 290 words allocated in the GSL, AWL, FWL, AL and PNL. Some examples of words eliminated were *towel, a, with, add, ice, classic, alternatively, Thai, for*, and *bowl.* In total, 524 potential words remained. Some examples of the remaining words were *allspice, balsamic, chopped, drizzle, fillet, garnish, marinate, minced, oregano*, and *simmer.* Some sample words with their contexts extracted from the FSC are provided below.

Thread 4: Allspice

*To make the coconut rice, heat the oil in a saucepan and cook the garlic, thyme, **allspice** and rice until fragrant.*

Thread 5: Drizzle

*This classic appetizer consists of a grilled slice of bread rubbed with garlic and **drizzled** with extra virgin olive oil.*

Thread 6: Minced

*Shrimps that are sauteed in a pan with **minced** garlic, lemon juice, paprika, and olive oil*

Expert Viewpoints

After removing 290 words from the list, the remaining 524 words were considered by the three experts who judged whether the words were suitable for inclusion in the FSWL. However, as the experts could find it tiresome to view a long list of words that could lead to inaccurate feedback (Laosrirattanachai & Ruangjaroon, 2021), before distributing the word-list check list to experts, the 524 words that had met the previous two criteria were transformed from word type into word family using the VocabProfile programme (Cobb, 2021). This shortened the word list to 347 word families. Some examples were *chop, chops*, and *chopped* which were transformed into *chop*, while *spice, spices, spiced*, and *spicy* were changed to *spice*. The expert's viewpoints were the last criterion applied to construct the FSWL. The results from the experts indicated that there were 261 words rated as 3 (words with a meaning very relevant to the food service field) or 4 (words with a meaning specific to the food service field and not being found in other areas) by at least two of the experts. These 261 verified words were included in the final version of the Food Service Word List. Some examples of words that met this third criterion were *horseradish, cauliflower, fry, ravioli, gnocchi, glaze, saucepan, shallot, seasoning*, and *zest.* Some examples of words rejected from the FSWL were *quarter* and *rectangle*.

Some examples of words included in the FSWL are illustrated with sentences extracted from the FSC as follows.

Thread 7: Fillet

Example 1: *It is made with raw fish **fillets** cut into bite-sized squares, stripes or thin slices, and marinated in lime or, rarely, lemon juice.*

Example 2: *Cut the pork open with a sharp knife and open the **fillet** out along its length, being careful not to cut all the way through the fillet.*

Example 3: *Dip the cod **fillets** into the batter and fry until golden-brown and cooked through.*

Thread 8: Simmer

Example 4: *Bring to the boil, then **simmer** until the sugar is a dark golden caramel.*

Example 5: *Add the butter and **simmer** gently for 3–4 minutes. Drain and keep warm.*

Example 6: *Add the sage and the brandy (or calvados) and **simmer** to burn off the alcohol.*

Thread 9: Garnish

Example 7: ***Garnish** with a whole basil leaf placed on top of the dressing.*

Example 8: *Serve the soup in bowls and **garnish** each portion with a poached egg.*

Example 9: ***Garnish** each serving with a drizzle of soured cream or a teaspoonful of Greek-style yoghurt.*

Thread 10: Balsamic

Example 10: *Add the **balsamic** vinegar and season with salt and pepper.*

Example 11: *Toss the dandelion leaves with **balsamic** and scatter on top.*

Example 12: *Add the **balsamic** vinegar and honey to the hot frying pan, and stir together until well combined.*

Thread 11: Griddle

Example 13: *The sandwich can be heated on a **griddle**, grilled, pan-fried, or toasted in a panini toaster.*

Example 14: *In Shanghai the wrapper is made from wheat flour dough briefly cooked on a **griddle**.*

Example 15: *Tortillas are cooked shortly on both sides on a cast-iron **griddle** comal, after which they are ready to become quesadillas.*

Corpus Coverage

As the developed FSWL has no comparable lists, it was compared with other word lists related or close to the food service field, specifically the Hotel Business Word Lists (Laosrirattanachai & Ruangjaroon, 2021) and the Composition Culinary Course Word List (Nordin et al., 2013). The results of the comparison are shown in Figure 3.

Figure 3

Food Service Corpus Coverage of Each Word List
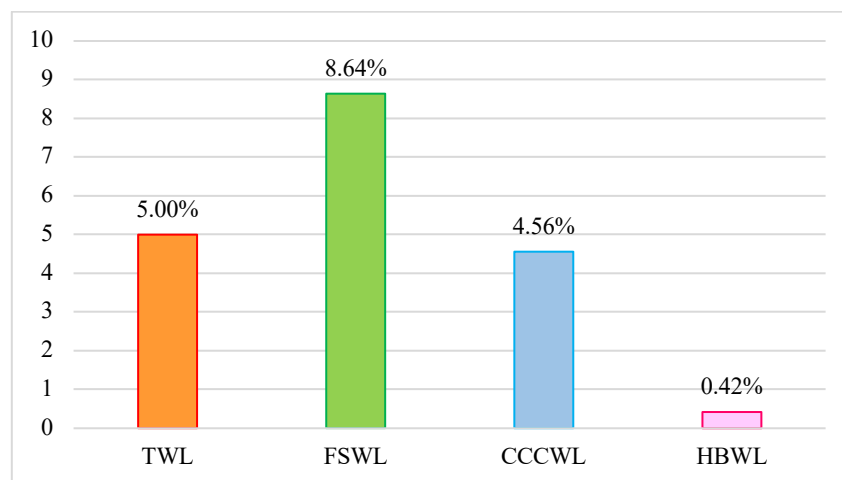
From Figure 3, a technical word list (TWL) is recommended to cover about 5% of the corpus. The Food Service Word List (FSWL) had the highest percentage coverage compared to the Composition Culinary Course Word List (CCCWL) with 4.56% and the Hotel Business Word List (HBWL) covering the corpus with only 0.42% of the FSC. This reflects that the FSWL is more effective and can satisfy learners in the food service industry.

Discussion

To evaluate a word list, a number of studies have used lexical coverage as the basis for comparison of the created word list with other related word lists (Dang & Webb, 2016). Using the corpus from which the word list was generated to evaluate the list is problematic and can be considered invalid (Coxhead, 2000). To avoid this issue, some studies evaluated the list using corpora different from those used to generate the list (Brezina & Gablasova, 2015; Gilner & Morales, 2008; Nation, 2004). The present study developed a corpus to evaluate the lexical coverage relevant to food service. However, there are no food service corpora available. Consequently, the FSC was used to evaluate the lexical coverage in the present study.

The GSL generally covers about 80% of each text (Nation & Waring, 1997), the AWL covers about 10% of each text (Coxhead, 2000), and created technical word lists should cover around 5% of each text. After the FSWL had been created, it was investigated to determine whether it covered 5% of

the FSC using AntWordProfiler programme, with the first 1,000 words in GSL, the second 1,000 words in GSL, AWL, AL, PNL and FSWL as the reference word lists. The result showed that the FSWL covered approximately 8.64% of the FSC. However, lexical coverage may differ by discipline (Chung & Nation, 2003; Coxhead et al., 2020; Hyland & Tse, 2007). Therefore, for the food service field, the FSWL as a technical word list covered more than expected 5% of the corpus. Learning the FSWL facilitates learners to increase their vocabulary range, and full understanding of the FSWL would suggest that learners would understand the technical vocabulary used in the food service industry.

Apart from the corpus coverage shown in the results section (8.64% for the FSWL, 4.56% for the CCCWL, and 0.42% for the HBWL), the sources used to construct the two word lists also strengthen the learning of the FSWL rather than the CCCWL and the HBWL. The CCCWL was compiled by gathering the PowerPoint slides from one writing course, and the list aimed to enhance writing ability. Therefore, the CCCWL might not be suitable for teaching in an ESP class in a food service industry course. The HBWL comprises specialised words used in the hotel business. Food service might be one of the sub-fields in the hotel; however, it is unlikely that the HBWL contains adequate specialised words and could facilitate learners to study food service in an ESP classroom. Thus, the lack of a specialised word list for food service emphasises the importance of creating the FSWL for the benefit of learners.

Limitations and Recommendations

While the four websites provided data concerning top foods from many nationalities worldwide, not all nationalities were included. This may result in incomplete coverage of all important words used in the food service field around the world. Some countries may use words that refer to specific ingredients, cooking methods, or special kitchenware that do not appear in the FSWL. Therefore, we recommend collecting a broader range of worldwide data from as many countries as possible to develop a better FSWL so that users of this word list can increase their vocabulary range and communicate more effectively. The present study limited itself to analysis only at the single-word level and used only the relative frequency. We recommend further study using the absolute frequency and comparing the list with the FSWL. Furthermore, a multiword unit list is suggested to strengthen the vocabulary of learners in the food service industry.

Pedagogical Implications

We can apply the FSWL in teaching and learning in various ways. When viewing the challenge of implementing a word list in teaching and learning settings, particularly with Thai learners, gaining learners' attention in using the list seems to be one of the most difficult tasks. Rarely are learners interested in studying the list, even after it has been assigned by a teacher. According to Zhu (2012), online social media in collaboration with teaching and learning significantly contributed to learners' academic performance and satisfaction. In addition, online social media increase the opportunities for

learners to access and share ideas, knowledge, and learning materials with each other (Cavus & Ibrahim, 2009; Gikas & Grant, 2013; Richardson & Lenarcic, 2008). Therefore, we suggest making instructional packages for learning on social media that can be easily accessed and are user-friendly for learners in the food service field. For example, the teacher can create food service vocabulary content with sample pictures and text extracted from the FSC and post them on a Facebook Fan Page, Instagram, Twitter, YouTube, and blog. In addition, teachers can create a textbook or teaching materials in class for English for Food Service or relevant courses using vocabulary in the FSWL.

Conclusion

A systematic procedure and the application of relevant criteria to construct the FSWL are necessary to create a useful Food Service Word List to facilitate and support learners or people who are related to or interested in the food service industry. The present study gathered data from four reliable websites related to food service to compile the Food Service Corpus. In total, the FSC contained 1,871,271 tokens. Then, three criteria (keyword analysis, lexical profiling, and expert viewpoints) were applied to generate the FSWL. First, keyword analysis was used to consider unusual frequency words occurring in the FSC compared with the British National Corpus based on their log-likelihood values using the Key-BNC programme. Then, the 814 words considered passing the keyword analysis criterion were analysed using the AntWordProfiler programme to develop a lexical profile. As a result, 290 words allocated in the GSL, AWL, FWL, AL and PNL were removed, leaving 524 words to be considered by the three experts under the next criterion. After shortening the 524 word types to 347 word families, the word-list checklists containing 347 words were scored using a 4-rating scale by the three experts in the field of food service industry. Finally, 261 words passed all criteria and were contained in the Food Service Word List.

References

Anthony, L. (2021). *AntWordProfiler* (Version 1.5.1) [Computer Software]. Tokyo, Japan: Waseda University. https://www.laurenceanthony.net/software.

Aroonmanakun, W. (2011). Corpora and emerging technology for ELT. *Journal of Studies in the English Language, 6*(1), 221–236.

Attachoo, B., & Chaturongkul, P. (2015). A study of vocabulary size, competency, learning strategies, and perceptions of problems with vocabulary learning of students with learning disabilities. *LEARN Journal: Language Education and Acquisition Research Network, 8*(2), 59–83.

Bennett, R. G. (2010). *Using corpora in the language learning classroom: Corpus linguistics for teachers*. University of Michigan Press.

Brezina, V., & Gablasova, D. (2015). Is there a core general vocabulary? Introducing the New General Service List. *Applied Linguistics, 36*(1), 1–22.

Browne, C. (2013). The new general service list: Celebrating 60 years of vocabulary learning. *The Language Teacher,* 4(37), 13–16.

Cavus, N., & Ibrahim, D. (2009). M-learning: An experiment in using SMS to support learning new English language words. *British Journal of Educational Technology, 40*(1), 78–91.

Chung, T., & Nation, I. S. P. (2003). Technical vocabulary in specialized texts. *Reading in a Foreign Language, 15*(2), 103–116.

Chung, T., & Nation, I. S. P. (2004). Identifying technical vocabulary. *System, 32*(2), 251–263.

Cobb, T. (2021). *Web Vocabprofile.* [Online programme]. http://www.lextutor.ca/vp/.

Coxhead, A. (2000). A new academic word list. *TESOL Quarterly, 34*(2), 213–238.

Coxhead, A., Parkinson, J., Mackay, J. & McLaughlin, E. (2020). *English for vocational purposes.* Routledge.

Crystal, D. (1991). *A dictionary of linguistics and phonetics.* Blackwell Publishing.

Dang, T. N. Y., & Webb, S. (2016). Evaluating lists of high-frequency words. *International Journal of Applied Linguistics, 167*(2), 132–158.

Food Intelligence Center. (2020). *Thailand Food Industry Profile 2019.* http://fic.nfi.or.th/info_graphic_detail.php?id=28.

Gabrielatos, C., & Marchi, A. (2012). Keyness: Appropriate metrics and practical issues. *Paper presented at the CADS International Conference*, Bologna, Italy.

Garcia, F. A. (2014). How language learners can benefit from corpora, or not. *Recherches en Didactique des Langues et des Cultures, 11*(1), 1–15.

Gardner, D., & Davies, M. (2014). A new academic vocabulary list. *Applied Linguistics, 35*(3), 305–327.

Gikas, J., & Grant, M. (2013). Mobile computing devices in higher education: Student perspectives on learning with cellphones, smartphones & social media. *Internet and Higher Education Mobile, 19*, 18–26

Gilner, L., & Morales, F. (2008). Corpus-based frequency profiling: Migration to a word list based on the British National Corpus. *The Buckingham Journal of Language and Linguistics, 1*, 41–58.

Graham, D. (2021). *Key-BNC.* [Software]. http://crs2.kmutt.ac.th/Key-BNC/.

Gries, S. T. (2015). Quantitative designs and statistical techniques. In D. Biber, & R. Reppen (Eds.), *The Cambridge Handbook of English corpus linguistics* (pp. 50–71). Cambridge University Press.

Hyland, K., & Tse, P. (2007). Is there an "Academic Vocabulary"? *TESOL Quarterly, 41*(2), 235–253.

Karim, S., & Chi, C. (2010). Culinary tourism as a destination attraction: an empirical examination of destinations' food image. *Journal of Hospitality Marketing & Management, 19*(6), 531–555.

Konstantakis, N. (2007). Creating a business word list for teaching business English. *Estudios De Linguistica Inglesa Aplicada (ELIA), 7*, 79–102.

Kruawong, T., & Phoocharoensil, S. (2020). Developing an English zoology academic word list: A corpus-based study. *Thoughts, 2*, 63–78.

Laosrirattanachai, P., & Laosrirattanachai, P. (2021). Applying lexical profiling to construct technical word lists for Thai tourist guides. *PASAA*, *62*(1), 61–91.

Laosrirattanachai, P., & Ruangjaroon, S. (2020). The word lists of hospitality service review construction. *Journal of Studies in the English Language, 15*(1), 107–158.

Laosrirattanachai, P., & Ruangjaroon, S. (2021). Corpus-based creation of tourism, hotel, and airline business word lists. *LEARN Journal: Language Education and Acquisition Research Network, 14*(1), 50–86.

Laufer, B. (1989). What percentage of text-lexis is essential for comprehension? In C. Laurén & M. Nordmann (Eds.), *Special language: From humans thinking to thinking machines* (pp. 316–323). Multilingual Matters.

Laufer, B., & Nation, I. S. P. (1995). Vocabulary size and use: Lexical richness in L2 written production. *Applied Linguistics, 16*(3), 307–332.

Lei, L., & Liu, D. (2016). A new medical academic word list: A corpus-based study with enhanced methodology. *Journal of English for Academic Purposes, 22*, 42–53.

Lestari, I. W., & Hardiyanti, N. (2020). Vocabulary learning autonomy through incorporation of English songs: Indonesian EFL students' perspectives. *3L: The Southeast Asian Journal of English Language Studies, 26*(2), 91–104.

Lindquist, H. (2009). *Corpus linguistics and the description of English*. Edinburgh University Press.

Liu, J., & Han, L. (2015). A corpus-based environmental academic word list building and its validity test. *English for Specific Purposes, 39*, 1–11.

Nordin, M. N. R., Stapa, S. H., & Darus, S. (2013). Developing a specialized vocabulary word list in a composition culinary course through lecture notes. *Advances in Language and Literary Studies, 4*(1), 78–88.

Martinez, I., Beck, S., & Panza, C. (2009). Academic vocabulary in agriculture research articles: A corpus-based study. *English for Specific Purposes, 28*(3), 183–198.

Muñoz, V. (2015). The vocabulary of agriculture semi-popularization articles in English: A corpus-based study. *English for Specific Purposes, 45*(3), 26–44.

Nation, I. S. P. (2004) A study of the most frequent word families in the British National Corpus. In P. Bogaards & B. Laufer (eds.) *Vocabulary in a Second Language: Selection, Acquisition and Testing Amsterdam* (pp. 3–13). John Benjamins.

Nation, I. S. P. (2006). How large a vocabulary is needed. *The Canadian Modern Language Review, 63*(1), 59–83.

Nation, I. S. P. (2013). *Learning vocabulary in another language.* Cambridge University Press.

Nation, I. S. P. (2016). *Making and Using Word Lists for Language Learning and Testing.* John Benjamins.

Nation, I. S. P. (2018, April, 10). *Resources.* https://www.wgtn.ac.nz/lals/resources.

Nation, I. S. P., & Waring, R. (1997). Vocabulary size, text coverage and word lists. In N. Schmitt & M. McCarthy (Eds.), *Vocabulary, Description, Acquisition and Pedagogy* (pp. 6–19). Cambridge University Press.

Nosratinia, M., Mirzakhani, E., & Zaker, A. (2013). Toward a humanistic instruction: Collaborative strategic reading approach and EFL learners' reading comprehension. *International Journal of Advanced Studies in Humanities and Social Science, 1*(8), 1119–1138.

O'Keeffe, A. & McCarthy, M. (2010). Historical perspective: What are corpora and how have they evolved?. In A. O'Keeffe, & M. McCarthy (eds), *The Routledge Handbook of Corpus Linguistics* (pp. 3–13). Routledge.

O'Keeffe, A., McCarthy, M., & Carter, R. (2007). *From corpus to classroom: Language use and language teaching*. Cambridge University Press.

Paakki, H. (2013). *Difficulties in speaking English and perceptions of accents - A Comparative Study of Finnish and Japanese Adult Learners of English.* M.A. Thesis, University of Eastern Finland.

Richardson, J., & Lenarcic, J. (2008). Text messaging as a catalyst for mobile student administration: The "trigger" experience. *International Journal of Emerging Technologies & Society, 6*(2), 140–155.

Schmitt, N. (2010). *Researching vocabulary: A vocabulary research manual*. Palgrave Macmillan.

Soonthornsaratool, Y. (2018). Regulating food trucks in Thailand. *Thammasat Business Law*, *8*, 143–156.

Tangpijaikul, M. (2014). Preparing business vocabulary for the ESP classroom. *RELC Journal, 45*(1), 51–65.

Valipouri, L., & Nassaji, H. (2013). A corpus-based study of academic vocabulary in chemistry research articles. *English for Specific Purposes, 12*(4), 248–263.

Ward, J. (2009). A basic engineering English word list for less proficient foundation engineering undergraduates. *English for Specific Purposes, 28*(3), 170–182.

Watson Todd, R. (2017). An opaque engineering word list: Which words should a teacher focus on? *English for Specific Purposes, 45*(1), 31–39.

West, M. (1953). *A general service list of English words*. Longman.

Wilkins, D. (1972). *Linguistics in language teaching.* Edward Arnold.

Yang, M. (2015). A nursing academic word list. *English for Specific Purposes, 37*(1), 27–38.

Zhu, C. (2012). Student satisfaction, performance, and knowledge construction in online collaborative learning. *Journal of Educational Technology & Society, 15*(1), 127–136.

Appendix

The 261 words of the Food Service Word List (FSWL)

1. allspice	2. almond	3. anchovy	4. anise	5. apricot
6. asparagus	7. aubergine	8. avocado	9. bacon	10. balsamic
11. barbecue	12. basil	13. basmati	14. beetroot	15. blanch
16. blend	17. bone	18. brandy	19. breadcrumbs	20. breast
21. broccoli	22. broth	23. bubble	24. bun	25. buttercream
26. buttermilk	27. butternut	28. cabbage	29. caper	30. carbohydrate
31. cardamom	32. casserole	33. caster	34. cauliflower	35. cayenne
36. celery	37. cheddar	38. chickpea	39. chill	40. chip
41. chives	42. chop	43. chorizo	44. chunk	45. chutney
46. cider	47. cinnamon	48. clam	49. cling	50. clove
51. coat	52. cocoa	53. cod	54. colander	55. coriander
56. cornflour	57. courgette	58. couscous	59. cream	60. crumble
61. crunch	62. cube	63. cumin	64. curry	65. custard
66. decorate	67. demerara	68. deseeded	69. dice	70. dill
71. dough	72. drain	73. drizzle	74. dumpling	75. fennel
76. feta	77. fillet	78. flake	79. flatbread	80. flatleaf
81. floret	82. flour	83. fluff	84. foil	85. freeze
86. fry	87. garlic	88. garnish	89. ghee	90. ginger
91. glaze	92. gnocchi	93. grate	94. gravy	95. griddle
96. grill	97. groundnut	98. haddock	99. halve	100. harissa
101. hazel	102. heatproof	103. homemade	104. horseradish	105. icing
106. ingredient	107. kale	108. ketchup	109. kimchi	110. knead
111. knob	112. ladle	113. lamb	114. lard	115. lardons
116. lasagne	117. leek	118. leftover	119. lemongrass	120. lentil
121. lettuce	122. lid	123. liqueur	124. marinade	125. marinate
126. masala	127. mascarpone	128. mash	129. mayonnaise	130. meatballs
131. meringue	132. mince	133. mint	134. mirin	135. miso
136. mortar	137. mozzarella	138. muffin	139. muscovado	140. mushroom
141. mussel	142. mustard	143. naan	144. nutmeg	145. olive
146. oregano	147. ounce	148. oven	149. ovenproof	150. pancake
151. pancetta	152. paneer	153. paprika	154. parchment	155. parmesan
156. parsley	157. passata	158. pasta	159. pastry	160. pea
161. peanut	162. pecorino	163. peel	164. pepper	165. peppercorns

166. pestle	167. pesto	168. pickle	169. pipe	170. pistachio
171. plum	172. pod	173. pomegranate	174. prawn	175. protein
176. puff	177. pumpkin	178. raisin	179. ramekin	180. rapeseed
181. raspberry	182. ravioli	183. recipe	184. rhubarb	185. ricotta
186. rinse	187. risotto	188. saffron	189. sage	190. salad
191. salsa	192. saucepan	193. sausage	194. sautã	195. scallop
196. seafood	197. seal	198. seasoning	199. semolina	200. sesame
201. shallot	202. shaoxing	203. sherry	204. shred	205. sieve
206. sift	207. simmer	208. skewer	209. skinless	210. slice
211. soak	212. sourdough	213. soy	214. spatula	215. spice
216. spinach	217. splash	218. sponge	219. squash	220. squeeze
221. squid	222. steak	223. stew	224. sticky	225. strain
226. sultana	227. sunflower	228. sushi	229. sweetcorn	230. syrup
231. tablespoon	232. taco	233. tagine	234. tamarind	235. tarragon
236. tart	237. teaspoon	238. thigh	239. thyme	240. toast
241. tofu	242. tortilla	243. toss	244. trim	245. turmeric
246. vanilla	247. vegan	248. vegetarian	249. venison	250. vinegar
251. virgin	252. walnut	253. watercress	254. wax	255. wedge
256. whisk	257. wilt	258. wok	259. yeast	260. yolk
261. zest				

About the Authors

Thanawan Rungrueang, **Pitchapha Boonprasert,** and **Sitanan Poempongsajaroen** are students in the English for Service Industry programme at the Faculty of Hospitality Industry, Kasetsart University, Nakhonpathom, Thailand.

Piyapong Laosrirattanachai is an Assistant Professor at the Faculty of Hospitality Industry, Kasetsart University, Thailand. He holds a PhD in the field of Linguistics from Srinakharinwirot University, Thailand. His interests include teaching and learning vocabulary, word list construction, corpus linguistics, and genre and move analysis.