CORPUS-BASED ANALYSIS OF ENGLISH VOCABULARY IN PHARMACEUTICAL LEAFLETS AND A DESIGN OF A SAMPLE DICTIONARY

PHATRINTHORN PONGNAN

A THESIS SUBMITTED IN PARTIAL FULFILLMENT
OF THE REQUIREMENT FOR THE DEGREE OF
MASTER OF ARTS IN APPLIED LINGUISTICS
(ENGLISH FOR SCIENCE AND TECHNOLOGY)
SCHOOL OF GRADUATE STUDIES
KING MONGKUT'S INSTITUTE OF TECHNOLOGY LADKRABANG
2006

ISBN 974-15-2752-7

CORPUS-BASED ANALYSIS OF ENGLISH VOCABULARY IN PHARMACEUTICAL LEAFLETS AND A DESIGN OF A SAMPLE DICTIONARY

PHATRINTHORN PONGNAN

A THESIS SUBMITTED IN PARTIAL FULFILLMENT

OF THE REQUIREMENT FOR THE DEGREE OF

MASTER OF ARTS IN APPLIED LINGUISTICS

(ENGLISH FOR SCIENCE AND TECHNOLOGY)

SCHOOL OF GRADUATE STUDIES

KING MONGKUT'S INSTITUTE OF TECHNOLOGY LADKRABANG

2006

ISBN 974-15-2752-7

หัวข้อวิทยานิพนธ์	การวิเคราะห์คำศัพท์จากคลังข้อมูลทางภาษาของเอกสารกำกับยาภาษาอังกฤษและการสร้างพจนานุกรมตัวอย่าง
นักศึกษา	นางสาวภัทรินธร ปองนาน
รหัสประจำตัว	4064373
ปริญญา	ศิลปศาสตรมหาบัณฑิต
สาขาวิชา	ภาษาศาสตร์ประยุกต์-ภาษาอังกฤษเพื่อวิทยาศาสตร์และเทคโนโลยี
พ.ศ.	2549
อาจารย์ผู้ควบคุมวิทยานิพนธ์	รศ. ดร. จิราภา วิทยาภิรักษ์
อาจารย์ผู้ควบคุมวิทยานิพนธ์ร่วม	ผศ. ดร. จิราภรณ์ อินทราไสย

บทคัดย่อ

การศึกษานี้มีจุดประสงค์เพื่อวิเคราะห์คำศัพท์ในเอกสารกำกับยาภาษาอังกฤษ เอกสารกำกับยาที่ใช้ในการวิจัยนี้เก็บจากเอกสารกำกับยา 300 ชิ้น ที่เกี่ยวข้องกับการยาที่ใช้รักษาอาการของโรคที่เกี่ยวข้องกับการป่วยและการตาย 9 กลุ่มโรค เครื่องมือที่ใช้ในการวิเคราะห์คำศัพท์ได้แก่ โปรแกรมจำแนกคำศัพท์ตามความถี่เวิร์ดสมิธ(Wordsmith Tools) รุ่นที่ 3.0 ค่าสถิติที่ใช้ในงานวิจัยคือ ค่าความถี่ และเปอร์เซ็นต์ คำทั้งหมดที่พบจากการศึกษาครั้งนี้แบ่งออกเป็น 4 กลุ่ม ได้แก่ 1) กลุ่มคำศัพท์พื้นฐานที่ตรงตามบัญชีคำศัพท์พื้นฐานซึ่งจัดทำโดย Bauman และ Culligan (1995) 2) กลุ่มคำศัพท์ทางวิชาการที่ตรงตามบัญชีคำศัพท์ทางวิชาการซึ่งรวบรวมโดย Coxhead (2000) 3) กลุ่มคำศัพท์เทคนิคซึ่งระบุได้จากพจนานุกรมแพทย์ศาสตร์ 4) กลุ่มคำอื่นๆที่ไม่สามารถจัดอยู่ใน 3 กลุ่มข้างต้น

ผลการวิเคราะห์คำศัพท์พบว่า ศัพท์ที่ใช้ในเอกสารกำกับยามีลักษณะเป็นคำศัพท์เทคนิคค่อนข้างสูง จากคำที่ศึกษาทั้งหมด 1,223 คำ ประกอบไปด้วยคำศัพท์พื้นฐานร้อยละ 38.68 คำศัพท์ทางวิชาการร้อยละ 9.24 คำศัพท์เทคนิคร้อยละ 33.03 นอกจากนี้ยังพบคำศัพท์อื่นๆที่ไม่ได้ปรากฏอยู่ในกลุ่มข้างต้นประมาณร้อยละ 19.05 ในส่วนของตัวย่อที่ใช้พบลักษณะของการย่อคำทั้งหมด 5 วิธี ได้แก่ การย่อแบบตัดท้ายคำ, การย่อแบบคงตัวอักษรแรกของคำเต็ม, คำย่อที่มาจากการผสมคำของอักษรต้น คำย่อแบบย่นคำ และคำย่อแบบแทนคำ

ผลที่ได้จากการศึกษาความถี่ของคำศัพท์ และ คอนคอร์แดนซ์ (concordance) ได้นำมาสร้างเป็นตัวอย่างคำที่บรรจุอยู่ในพจนานุกรมตัวอย่าง ซึ่งประกอบด้วย ความหมายทั้งภาษาไทยและอังกฤษ การออกเสียงคำ ไวยากรณ์ รากศัพท์ ตัวอย่างการใช้ และภาพประกอบ เพื่อใช้ประโยชน์ในการสร้างพจนานุกรมตัวอย่างสำหรับนักศึกษาและบุคลากรที่เกี่ยวข้องกับสาขาแพทยศาสตร์

Thesis Title	Corpus-based Analysis of English Vocabulary in Pharmaceutical Leaflets and a Design of a Sample Dictionary
Student	Miss Phatrinthorn Pongnan
Student ID.	44064373
Degree	Master of Arts
Program	Applied Linguistics-English for Science and Technology
Year	2006
Thesis Advisor	Assoc. Prof. Dr. Jirapa Vitayapirak
Thesis Co-Advisor	Asst. Prof. Dr. Jiraporn Intrasai

ABSTRACT

The purpose of this study was to analyze vocabulary used in English pharmaceutical leaflets. Three hundred of pharmaceutical leaflets analyzed in the corpus were collected from the leaflets which used for curing or treating the disease of the nine groups of cause of illness and cause of death. The analysis tool in this study was the concordancing software *'Wordsmith Tools Version 3.0'*. The statistical devices employed in this corpus study were frequencies and percentages. All words in this study were divided into three groups. The first group comprised general vocabularies that appeared in the General Service List (GSL) created by Bauman and Culligan (1995). The second group comprised the academic vocabularies that appeared in the Academic Word List (AWL) compiled by Coxhead (2000). The third group comprised technical vocabularies that identified by medical dictionaries. The forth group is the word outside those three groups.

The results of this study showed that the vocabulary in pharmaceutical leaflets was highly technical. According to the corpus findings, the ten most frequent words were function words. Regarding the words in this study, there were 1,223 lemmas which consisted of 38.68 % of the general vocabularies, 9.24 % of the academic vocabularies, 33.03 % of the technical vocabulary, and 19.05 % of the words outside those three groups. In terms of abbreviations, there were five main types: clippings, initialisms, acronyms, contractions, and substitutions.

The corpus findings in terms of word frequencies and concordance are used to develop model entries for the sample dictionary of medicine. It offers bilingual definition, pronunciation, grammatical information, root, example of usage and illustration. This study can be beneficial for the learners and medicine-related people.

ACKNOWLEDGEMENT

I would like to express my deepest gratitude to my thesis advisor Associate Professor Dr. Jirapa Vitayapirak and Co-advisor Assistant Professor Dr. Jiraporn Intrasai for their valuable advice, devotion, insightful comments and suggestions. Their encouragement and knowledge of how to conduct the research are really valuable. Without them, I could not reach the goal.

I wish to express my thanks to the School of Graduate Studies, King Mongkut's Institute of Technology Ladkrabang (KMITL) for giving me a scholarship to carry out this thesis.

Finally, my appreciation goes to all of my friends for their sympathy and encouragement throughout the years. Most of all I would like to give special appreciation to my parents and sister for their love and care. A lot of thanks are also for my relatives for their support.

Phatrinthorn Pongnan

TABLE OF CONTENTS

TABLE OF CONTENTS (continued)

TABLE OF CONTENTS (continued)

LIST OF TABLES

LIST OF FIGURES

CHAPTER 1

INTRODUCTION

1.1 Statement and Significance of the Problem

At present, science and technology have brought advances in every area of medicine. Despite the proliferation of medical science, medical scientists know that the search for better medical care will never end. As human beings alter their life styles and environments, new diseases and debilitating conditions develop to challenge medical research.

It is generally acknowledged that most new medical innovations are from English speaking countries. English is used to present ideas in medical science. Medical scientists, both native and non-native speakers of English, write research documents in the English language for promoting their work worldwide. This has led to an increasing number of English medical articles and related papers being published throughout the world. Therefore, it is essential for Thai medical students to study English for Medical Purposes which they will use in their future career.

In Thailand, it is widely recognized that medical science as one of the most wanted courses at university educational level, because of the increasing need for skilled medical doctors in the job markets. This has imposed upon language teachers a pressing need to establish what language features and skills are most needed by medical doctors in order to design relevant and practical programs to prepare medical science students for the job market.

There has been a substantial amount of research on the study of medical language. A study by Prawanpath (1984) found that English is necessary for the medical professional in the Thai environment. Pre-medical science students at Mahidol University have a good general English vocabulary in a daily-life context. However, they lacked an experience in using medical terms at the university level. Methold (1997: vi) stated that medical writing relies very heavily on a specialized vocabulary with a great deal of vocabulary in English and Greek. English medical terminology is an important characteristic of the language of medicine. It is usually used to explain syndromes, lesions, surgical procedures that occur either in textbooks or in clinical usage. As a consequence, it is becoming more widely accepted that the knowledge of medical lexical items in English is needed to

integrate with the medical curriculum because medical English terminology is an essential element of medical study and practice.

It can be seen from the above situations that an analysis of English vocabulary in medical science can be beneficial for Thai students. Therefore, this thesis attempts to use the corpus of the pharmaceutical leaflets in finding out the medical terminology. This study can be useful for medicine-related people such as medical, pharmacy and nurse students, linguists who are interested in vocabulary used in English for Medical Purposes (EMP), and patients or people who use drugs and want to make better understanding of English used in medical leaflets.

1.2 Goals and Objectives

The purpose of this study is to analyze the English vocabulary based on the pharmaceutical corpus. The present study has eight objectives as follows:

1.2.1 To analyze general vocabulary in terms of word lists.

1.2.2 To classify the general vocabulary into nouns, verbs, adjectives, and adverbs.

1.2.3 To analyze academic vocabulary in terms of word lists.

1.2.4 To classify the academic vocabulary into nouns, verbs, adjectives, and adverbs.

1.2.5 To analyze technical vocabulary in terms of word lists.

1.2.6 To classify the technical vocabulary into nouns, verbs, adjectives, and adverbs.

1.2.7 To identify abbreviations and collocations.

1.2.8 To design sample dictionary entries based on the technical vocabularies found in the corpus.

1.3 Scope or Limitations of the Study

1.3.1 The vocabulary inputs of the study were taken from three hundred pharmaceutical leaflets published during 1997-2001 and they were collected from three hospitals in Bangkok, namely Chulalongkorn, Siriraj and Ramathibodi Hospitals.

1.3.2 The pharmaceutical leaflets were collected from the drugs related to the first five causes of illness and first five causes of death related to Thai people according to the Public Health statistics. (Ministry of Public Health 2000). The first five causes of

illness were respiratory system diseases, digestive system diseases, musculoskeletal system and connective tissue diseases, skin and subcutaneous tissue system diseases, and circulatory system diseases. The first five causes of death were malignant neoplasm and all forms; circulatory system diseases; pneumonia and other diseases of lung; nephritis, nephrotic syndrome and nephrosis; and disease of liver and pancreases.

1.3.3 The *General Service List* (Bauman and Culligan 1995) was used as the guidelines for identifying the general vocabulary.

1.3.4 The *Academic Word List* (Coxhead 2000) was used as the guidelines for identifying the academic vocabulary.

1.3.5 The *Mosby's Medical, Nursing, & Allied Health Dictionary* (Anderson et al. 2002) and *Merriam Webster: Medical Dictionary Online* were used as the guidelines for identifying the technical vocabulary.

1.4 Process of the Study

The process of the study is as follows:

1.4.1 Literature Review

The literature and related research concerning vocabulary, corpus-based approach and lexicography were studied.

1.4.2 Data Collection

1) Three hundred pharmaceutical leaflets were collected on the basis of the first five causes of illness and the first five causes of death.

2) In the corpus compilation, all printed pharmaceutical leaflets were scanned and stored as plain-text files (*.txt).

1.4.3 Data Analysis

1) The computer software program, WordSmith Tools Version 3.0 was used to examine tokens and word types and calculate type/token ratio and word frequency of the whole corpus. All

texts were categorized into the groups of vocabulary, namely general, academic, and technical vocabulary. Then they were presented in terms of percentage of word occurrence.

2) From the word lists, the identification of general vocabulary is based on the General Service List provided by Bauman and Culligan (1995).

3) The general vocabulary was classified into nouns, verbs, adjectives, and adverbs.

4) From the word lists, the identification of academic vocabulary is based on the Academic Word List provided by Coxhead (2000).

5) The academic vocabulary was classified into nouns, verbs, adjectives, and adverbs.

6) From the word lists, the identification of technical vocabulary is based on The *Mosby's Medical, Nursing, & Allied Health Dictionary* (2002) and *Merriam - Webster Medical Dictionary Online* (2006).

7) The technical vocabulary was classified into nouns, verbs, adjectives, and adverbs.

8) The collocations were classified into lexical and grammatical collocations.

9) Based on the study of Barnhart et al. (1995), the abbreviations were examined.

1.4.3 Results and Discussion

The results were presented as follows:

1) The overall results were presented in terms of number of tokens and word types, type/token ratio, and word frequency list.

2) The frequency and percentage of occurrence of the general, academic, and technical vocabulary were presented.

3) The classification of general, academic, and technical vocabulary was presented in terms of nouns, verbs, adjectives, and adverbs.

4) Abbreviations and collocations were listed.

5) The sample dictionary entries were designed.

6) Discussion was provided.

1.4.4 Conclusions and Recommendations

Conclusions and recommendations for further studies were given.

1.5 Definition

Concordance a list of all the words which are used in a particular text or in the works of a particular author, together with a list of the contexts in which each word occurs. For example:

is 1 cap 3 times daily.	**Treatment**	may be conti
without Minirin	**treatment.**	A restricted w
nths. Children:	**Treatment**	Up to 2 year
1.5-6 mcg/ml. In the	**treatment**	of ventricular a
-15 mg/kg body	**Treatment**	of biliary reflu
g/day during long-	**treatment)**	Both lactab
0 mmHg or less)	**treatment**	is started or d
Patients assigned to	**treatment**	with the contr

(Concordance data from the pharmaceutical corpus)

Corpus (plural: corpora) a collection of text, of the written or spoken word, which is stored and processed on computer for the purposes of linguistic research (Renouf 1987: 1).

Content words the words which refer to a thing, quality, state, or action and which have meaning (lexical meaning) when the words are used alone. Content words are mainly nouns, verbs, adverbs, and adjectives, e.g. '*dose*', *'use'*, *'medical'*, *'quickly'*.

Function words the words which have a little meaning of their own, but which shows grammatical relationship in and sentences. Conjunctions, prepositions, articles, e.g. *'an'*, *'to'*, *'the'* are function words.

Lemma (plural: Lemmas) the word without all inflexional and derivational affixes. For example, the lemma of the words *'work'*, *'worked'*, *'working'* is *'work'*.

Lemmatization the process or result of dividing a text into lemma.

Word frequency lists	the list of vocabulary items in a corpus of language. This information can be shown in ether alphabetical order or frequency order.
The Academic Word List (AWL)	the list of academic words which are selects for tertiary level students who need to read and write academic prose. The AWL contains 570 word families based on a 3,500,000 token corpus of academic English.
The General Service List (GSL)	the most frequent 2,000 words of English (West 1953, Bauman and Culligan 1995).
Tokens or running words	the number of all words encountered in the text. Every word form is counted even though the same word form occurs more than once. For example, '*This medicine is not intended for women. Ask your doctor or pharmacist for advice before taking any medicine*'. These sentences contain eighteen tokens.
Vocabulary	A set of lexical items (words) in a language, also called the lexicon.
Word types	a token which occur more than once in different position was counted as one word type. For example, '*This **medicine** is not intended **for** women. Ask your doctor or pharmacist **for** advice before taking any **medicine**'*. These sentences have eighteen tokens but sixteen word types, as '***medicine***' and ***'for'*** occurs twice.

CHAPTER 2

LITERATURE REVIEW

This study intends to explore the English vocabulary in pharmaceutical leaflets. This chapter begins with the information concerning English for Specific purposes (ESP), English for Medical Purposes, previous research on medical English and background of linguistic corpora. Secondly, the vocabulary, word classes, collocations and abbreviations are provided. Finally, lexicography and technical dictionaries are reviewed.

2.1 English for Specific Purposes (ESP)

In the 1960s and early 1970s, the applied linguists gave special attention to describe the 'special' linguistic features of technical text. The word 'register' was used in association with the 'special language,' such as the language of biology, or the language of medicine. Special languages were analyzed for the frequency of occurrence of words in the texts and the particular syntactic and/or certain lexical items. The aim of the analysis was to identify the grammatical and lexical features of these registers and take them as their syllabus. The teaching materials then were based on descriptions of language use (Swales 1971).

ESP concentrates on language in context. It covers subjects varying from accounting or computer science to tourism and business management. As a matter of fact, ESP combines subject matter and English language teaching. Students are able to apply what they learn in their English classes to their main field of study, whether it is accounting, or medicine. Being able to use the vocabulary and structure that they learn in a meaningful context reinforces what is taught and increases their motivation.

ESP is an approach to language teaching in which all decisions as to content and method are based on the learners' reason for learning (Hutchinson and Waters 1987: 19). Richard et al. (1992: 124) has defined 'English for Specific Purposes' (ESP) in Dictionary of Language Teaching & Applied Linguistics:

> *'the role of English in a language course or program of instruction in which the content and aims of the course are fixed by the specific need of particular group or learners. For example course for English for Academic Purposes, English for Science and Technology, teaching, learning or studying English for Specific (e.g. scientific or technical) Purposes.'*
>
> (Richard et al. 1992: 124)

It is possible to distinguish ESP courses by the nature of the students' specialism. There has always been a purpose behind language teaching, and one of the more valuable aspects of ESP is that it has concentrated the mind on the ends a person want to pursue. Teaching English for a special purpose has to be clear about what purpose is, and to consider the ways to achieve this end (Kerr 1979:11). Mackey and Mountford (1987: 6) proposed three kinds of purposes:

1. *Occupational requirements*, e.g. for international telephone operators, civil airline pilots.
2. *Vocational training programmes*, e.g. for hostel and catering staff, technical trades.
3. *Academic or professional study*, e.g. for engineering, medicine, law.

This classification of purposes behind language teaching helps to define ESP varieties. It indicates three main types of ESP, differentiated according to whether the leaner requires English for 'Occupational', 'Vocational', or 'Academic' purposes. ESP courses were thus divided according to the purposes of the study, for example English for Tourism, English for Secretaries, English for Economics, or English for Medical Purposes.

2.2 English for Medical Purposes (EMP)

English for Medical Purposes (EMP) aims to meet the specific English language needs of the medical learners, for example, doctors, nurses, general physicians, dentists and other medical personnel. EMP focuses on themes and topics specific to the medical field and on a restricted range of the skills which may be required by medical learners such as writing medical reports and preparing a talk for medical meeting. All of these, certainly, involve a utilitarian purpose of the learners which

leads to the successful performance of work or the optimum effectiveness of medical training (Maher 1986: 112)

Pratt (2002) of the Faculty of Medicine, University of British Columbia suggested that English for Medical Purposes could not be taught at the level of basic English language teaching. He proposed that career-specific, highly technical language must be contextually based to ensure effectiveness in medical study and practice.

When the curriculum designer begins to develop a course in Medical English, he or she considers who the students are and what their motivations will be. Maher and Arbor (1992: 269-271) wrote *International Medical Communication in English* which was a good example of English for Medical Purposes. Their medical English guidance was divided into four broad categories:

1. *Speaking and listening* concern techniques of discussion at medical-meetings, seminar discussion, presenting papers at medical conferences, and telephoning in a hospital.
2. *Writing* involves writing reports, letters, notes, curriculum vitae, and scientific papers for publication. Sentences and structures are provided as well as a comprehensive selection of the most important types of the language function used in medical writing, advice to avoid sexism in written discourse and to spell correctly words which may cause confusion.
3. *Doctor-patient conversation/consultation* is concerned with taking a history, understanding symptoms, and the everyday language of doctor and patient communication.
4. *Miscellaneous* is concerned with medical terminology, medical abbreviations and information about the various English language examinations for medical graduates.

2.3 Previous Research on Medical English

This section is devoted to a review of research into the medical English. The earliest research to examine medical English was that of Bogart et al. (1995). In Belgium, they studied technical medical terms from medication package of full information leaflets in the member states of the European Community. Their aim was to select 1,400 scientific medical terms frequently used in the context of written drug information but difficult to understand for a lay person and create a glossary

in nine European languages (English, French, German, Dutch, Spanish, Portuguese, Italian, Greek and Danish). Their study was based on a corpus of Dutch scientific data sheets on medical products. Then the basic list of 1,830 medical terms stemed from a Dutch frequency search. This list was translated into the nine community languages as mentioned.

Later, Jabbour (1998) designed a research on 'corpus of medical research articles'. It was found that high frequency word, research word, and the medical terminology were used. Moreover, the analysis showed that present related tenses were related to writer's domain. Patterns of the present perfect, passive and active related to text attribution.

Paiva (1999) investigated style in a corpus of pharmaceutical leaflets. He compiled a corpus of 340 patient information leaflets from 40 companies. The aim was to obtain the variables that represented the main stylistic variation occurring in the texts. The findings were that the amount of linguistic variation in his corpus was more restricted than two large general corpora (LOB, London-Lund) because his corpus was a very specialized one.

2.4 Background of Linguistic Corpora

2.4.1 The Definition of Corpus

A corpus is any body of text with the aim of analysing its features (Landau 2001: 273). Nowadays the term 'corpus' (plural: corpora) refers to a collection of text, of the written or spoken word, which is stored and processed on computer for the purposes of linguistic research (Renouf 1987: 1).

Before having the electronic corpus, Randolph Quirk of University College London had begun his *Survey of English Usage*, although it was not done by computers and still exists in the form of endless trays of 4 x 6 inch paper slips stored in file cabinets.

The first major, computer-based study designed to be representative was the *Standard Corpus of Present-Day Edited American English* (better known as "*the Brown Corpus*"), assembled by W. Nelson Francis and Henry Kucera in 1963-64 at Brown University. It was destined to become a landmark in the development of corpus linguistics. The Brown Corpus paved the way for similar studies of British English and for much larger corpora in the decades that followed (Landau 2001: 278).

Modeled on the Brown Corpus, a study of British English was undertaken by Geoffrey Leech in 1970 at the University of Lancaster and completed in 1978 by the University of Oslo in collaboration with the Norwegian Computing Centre in Bergen; it has come to be known as the LOB Corpus (for Lancaster-Oslo/Bergen).

During the first half of the 19th century, the early corpora were produced to help language teachers trying to teach English systematically, and to simplify English for their learners (Jeffery 1953: v). In the field of ESL and EFL, language learners' corpora have been developed by various institutions, including Longman and Cambridge University Press. However, working on large corpora consumed a large amount of time, as electronic or computer tools were not available. Frequency lists such as Thorndike and Lorge's List (1944) in the United States and West's General Service List of English words (1953) in Britain were both derived from manually compiled corpora. Studies of word frequency and word frequency lists have become more interesting and challenging to English teachers and learners, with the development of the computer and information technology.

2.4.2 Types of Corpora

A corpus is complied for a particular purpose, and the types of corpus will depend on its objective. Corpus linguistics attempting to describe particular types of corpora have used a number of different terms to refer to them. In order to study features of the language in general, independently of the styles of particular types of text, it is necessary to use a *general corpus* or a collection of texts of as many different types as possible. In order to investigate the linguistic features that characterize a particular type of text, such as modern short stories in general, or newspaper, reports, or advertisement, it is necessary to use a corpus consisting of several examples of the appropriate type which is called a *specialized corpus* (Tribble and Jones 1990: 15). Hunston (2002: 14) stated that a corpus is always designed for a particular purpose, and the type of corpus will depend on it purpose.

2.4.2.1 General corpus

A general corpus is a collection of texts of as many types as possible. It may include written or spoken language, or both and unlikely to be representative of any particular 'whole', but will include as wide a spread of text as possible. A general corpus is usually much larger than a specialized corpus. It may be used to produce reference materials for language learning or translation,

and it is often used as a base-line in comparison with more specialized corpora. Because of this second function it is also called a *reference corpus*. Some well-known general corpora are such as the Brown Corpus, the LOB corpus, COBUILD.

2.4.2.1.1 The Brown Corpus

The first modern, electronically readable, corpus was the Brown Corpus of Standard American English. The corpus consists of one million words of American English texts printed in 1961. To make the corpus a good standard reference, the texts were sampled in different proportions from 15 different text categories: Press (repotage, editorial, reviews), Skills and Hobbies, Religious, Fiction (various subcategories), etc.

2.4.2.1.2 The LOB Corpus

The Lancaster-Oslo/Bergen Corpus (LOB) was compiled by researchers in Lancaster, Oslo and Bergen. The LOB corpus (British English) consists of 1 million words of written language, (500 texts of 2,000 words each) sampled in the same 15 categories as the Brown Corpus. For a long time, the Brown and LOB corpora were the only easily available computer readable corpora. Much research within the field of corpus linguistics has therefore been based on these corpora.

2.4.2.1.3 The COBUILD

In 1980, the compilers of the Collins COBUILD English Language Dictionary started to collect a corpus of texts on computer for dictionary making and language study. They had daily access to a corpus of approximately 20 million words. New texts were added to the corpus, and in 1991 it was launched as the Bank of English (BoE). More and more data have been added to the BoE, and the latest release (1996) contained some 320 million words.

2.4.2.2 Specialized corpus

A specialized corpus is a collection of texts of a particular type, such as newspaper editorials, geography textbooks, academic articles in a particular subject, lectures, a matter in medical leaflets, and so on. It is used to investigate a particular type of language. Researchers collect their own specialized corpora to reflect the kind of language they want to investigate.

2.5 Vocabulary

Vocabulary is a set of lexical items (words) in English (Crystal 1992: 412). There are two main kinds of word list: the General Service List (GSL), and the Academic Word List (AWL).

2.5.1 The General Service List (GSL)

In 1953 Michael West (West 1953) published his well-known General Service List of English words. This was a list of the 2000 most useful word families of English. The GSL is based on a 5 million word written corpus, excluding months, days of the week and numbers (see Appendix A). Each word in the list is accompanied by its inflected forms together with a list of common derivatives and compounds with the frequency and percentage of each major meaning of each word indicated. An updated version of the GSL was published by John Bauman and Brent Culligan in 1995. The list contains 2,284 words and presents in frequency order based on the Brown Corpus. Although the list has been criticized for many reasons, research into academic texts by Averil Coxhead (Coxhead 2000) has shown that the GSL covers almost 80% of the academic texts she studies. It would seem essential for any student to know these word families.

2.5.2 The Academic Word List (AWL)

Coxhead studied over 3.5 million words of academic text from 28 subject-areas and identified another 570 word families that were commonly used in academic texts from all subjects. The Academic Word List (AWL) of Coxhead was developed from a corpus of written academic texts by examining the range, frequency and uniformity of occurrence of words outside the first 2,000 words of English (West 1953). This corpus contains four disciplines (arts, commerce, law and science) and each discipline is made up of seven subject areas (see Table 2.1).

Table 2.1 Subject Areas in the Academic Corpus

Arts	Commerce	Law	Science
Education	Accounting	Constitutional Law	Biology
History	Economics	Criminal Law	Chemistry
Linguistics	Finance	Family Law and Medico-Legal	Computer Science
Philosophy	Industrial Relations	International Law	Geography
Politics	Management	Pure Commercial Law	Geology

Table 2.1 (continued)

Arts	Commerce	Law	Science
Psychology	Marketing	Quasi-Commercial Law	Mathematics
Sociology	Public Policy	Rights and Remedies	Physics

The AWL contains 570 words which have both high frequency and high coverage in academic texts, irrespective of subject area and discipline and accounts for approximately 10% of the total words (tokens) in the corpus described above. The words in AWL are ones which university students of English must be thoroughly familiar with for both reading and writing academic prose. It contains 570 words which have both high frequency and high coverage in academic texts, irrespective of subject area and discipline. These 570 words are not just word-forms, but consist of headwords plus their inflected and derived forms. Her words cover 10 % of the academic texts she studied. She called these words the Academic Word List (AWL). It would also necessary for any English for Academic Purposes students to know these words.

Therefore the General Service List plus the Academic Word List cover nearly 90% of the academic texts. English for Speicfic Purposes students would also need to know the specific word related to their subjects-around 5% of the words in an academic text and some of the less frequent words used in English about 5% (Nation 2001: 12).

2.5.3 Subject Specific Vocabulary

There are two other types of words students need to know, i.e. general words and the academic words. They also need to know the specific words for subject of learning. The subject specific words make up about 5% of the words of most academic texts. These words are very important to know and are usually learner in subject classes. They are used by specialists and are not generally known by non-specialists. In applied linguistics, ESP linguists also focus on the specialist vocabulary found in specific texts because one of the distinguishing features of ESP is the high concentration of terms.

Applied linguists such as Trimble and Trimble (1978), Robinson (1991), Dudley-Evans and St John (1998) classify vocabularies into three categories: '*subject specific vocabulary*', '*non subject-specific specialized vocabulary,*' and '*general vocabulary*'. The first group or subject

specific vocabulary used in one domain. They are monosemous with restricted, protected, and standardized meaning because one-to-one correspondence reduces ambiguity and improves communication such as *pharmacology*, *pediatric*, *geriatric*, and so on (Pearson 1998: 15). The second group or non subject-specific specialized vocabulary that has been identified is called *'semi-technical'*. It is comprises of words which occur in a number of scientific or technical areas such as, *absolute*, *accuracy*, *factor* and etc. The third group is *'general vocabulary'*.

In this study, words are categorized into three groups 1) general vocabulary, 2) academic vocabulary, and 3) technical vocabulary. The classification of the general vocabulary is by using the General Service List (Bauman and Culligan 1995), and academic vocabulary using the Academic Word List (Coxhead 2000). Medical and pharmaceutical dictionaries and the meaning in the context are used to classify the technical vocabulary.

2.6 Word Classes

Biber et al. (1999) categorize word classes into three classes according to their main functions and their grammatical behavior, i.e. lexical words, function words, and inserts.

2.6.1 Lexical words

The lexical words are the main carries of meaning in the text. There are four main classes of lexical words: nouns, verbs, adjectives, and adverbs.

2.6.1.1 Nouns

Noun typically refers to people, animals, places, things, or abstractions which can occur as the subject or object of a verb or object (complement) of a preposition. Nouns can be modified by an adjective and can be used with determiners e.g. acid, acne, dose.

2.6.1.2 Verbs

Verbs can have one of two major roles in verb phrases: main verb or auxiliary verb. Main verbs, such as the verb *went* in the following example, can stand alone as the entire verb phrase:

I went into the empty house.

In contrast, auxiliary verbs, such as the verbs *can* and *be* in the following example, occur together with some main verb (in this case *cited*):

Instances can be cited where this appears not to be the case.

There are three major classes of verbs: **lexical verbs** (also called full verbs, e.g. run, eat), **primary verbs** (be, have, and do), and **modal verbs** (e.g. can, will, might). These classes are distinguished by their roles as main verbs and auxiliary verbs.

2.6.1.3 Adjectives

Adjective is a word describing the thing, quality, state, or action which a noun refers to. For example, *side* in *the side effects* is an adjective. In English, adjectives usually have the following properties:

a. they can be used before a noun, e.g. a heavy bag
b. they can be used after *be, become, seem,* etc. as the bag is heavy.
c. they can be used after a noun as a complement, e.g. these books make the bag heavy.
d. they can be modified by an adverb, e.g. a very heavy bag.
e. they can be used in a comparative or superlative form, e.g. the bag seems heavier now.

2.6.1.4 Adverbs

Adverb is a word describing or adding to the meaning of a verb, an adjective, another adverb, or a sentence which answers such questions *as how ?, where?, or when?*. In English many adverbs have an *–ly* ending. For example, **adverbs of manner** e.g. *carefully, slowly*, **adverbs of place** e.g. *here, there, locally*, and **adverbs of time** e.g. *now, hourly, yesterday*.

2.6.2 Function words

While lexical words are the main building blocks of texts, **function words** provide the mortar which binds the texts together. The function words in this study are conjunctions, and prepositions.

2.6.2.1 Conjunctions

A conjunction is a word that connects individual words or groups of words. In fact, the word conjunction literally means "the act of joining" or "combination." There are three kinds of conjunctions: ***coordinating, correlative,*** and ***subordinating conjunctions.***

1. ***coordinating conjunctions*** connect words or groups of words that perform the same function in a sentence. The coordinating conjunctions are in the following list:

   and but for nor or yet

2. ***correlative conjunctions*** consist of two or more words that work together as a set. Correlative conjunctions function like coordinating conjunctions because they connect words that perform equal functions in a sentence. The correlative conjunctions are in the following list:

   either...or whether...or neither...nor

   not only...but (also) both...and

3. ***subordinating conjunctions*** connect subordinate clauses to independent clauses, which can stand by themselves as complete sentences. Subordinating conjunctions usually express relationships of time, manner, cause, comparison, or purpose.

Time	*after, as, as long as, as soon as, before, since, until, when, whenever, while*
Manner	*as, as if, as though*
Cause	*because*
Condition	*as long as, even if, if, provided that, unless*

2.6.2.2 Prepositions

Preposition is a word used with nouns, pronouns and gerunds to link them grammatically to other words. Most common prepositions are short, invariable forms: *as, at, by, down, for, from, in, into, on.*

2.6.3 Inserts

Inserts are a relatively newly recognized category of word. They do not form an integral part of a syntactic structure, but they are inserted rather freely in the text. They are often marked off by

intonation, pauses, or by punctuation marks in writing. The characteristic of the inserts are carried emotional and interactional meanings and are especially frequent in spoken texts such as *hm, uhhuh, ugh, yeah.*

In this thesis, the inserts are not studied because the inserts do not occur in the pharmaceutical leaflets.

2.7 Collocations

Collocation is the way in which words are used together regularly. Collocation refers to the restrictions on how words can be used together, for example which prepositions are used with particular verbs, or which verbs and nouns are together. For example, in English the verb ***perform*** is used with ***operation***, but not with ***discussion***:

The doctor performed the operation.

* *The committee performed a discussion.*

instead we say: *The committee held/had a discussion.*

The word ***'performe'*** is used with (collocates with) ***operation***, and ***hold*** and ***have*** collocate with ***discussion*** (Richard 1999: 62).

Collocations can be defined in numerous ways but for pedagogical purposes it is more practical to restrict the term to the following: two or three word clusters which occur repeatedly throughout spoken and written English. Bahns (1993) differentiated grammatical from lexical collocations as follows:

> *Examples of* ***grammatical collocations*** *include: account for, advantage over, adjacent to, by accident, to be afraid that... They consist of a noun, an adjective, or a verb, plus a preposition or grammatical structure such as an infinitive or clause.* ***Lexical collocations*** *on the other hand, do not contain prepositions, infinitives or clauses, but consist of various combinations of nouns, adjectives, verbs and adverbs.*
>
> (Bahns 1993:57)

2.8 Abbreviation

Abbreviation is a short way of writing a word or phrase, made by leaving out some of letters or by using only the first letter of each word. There are many abbreviations used in pharmaceutical leaflets. Those abbreviations are formed in different types according to the study of Barnhart (1995: xiii-xvi).

2.8.1 Clipping

The mechanism for clipping operates by omitting some part or parts at the end of a word, i.e. *strep* (for streptococcus), *auto* (for automobile), *dorm* (for dormitory), *condo* (for condominium), *demo* (for demonstration). Some clippings retain only the initial syllable of a word such as *pop* (for popular.

2.8.2 Apheresis

The mechanism for apheresis, which is sometimes considered a special type of clipping, omits some parts of the beginning of a word. As in clippings, many words formed by apheresis are originally considered to be nonstandard usage i.e. *bus* (for omnibus), *phone* (for telephone). The endings of ordinal numbers, e.g. *1st (for first), 2nd (for second), 3rd (for third), 4th (for fourth)* are included in aphersis.

2.8.3 Initialism

The process of grouping initial letters is the most familiar to the written abbreviation, i.e. *LA* (for Los Angles), *VCR* (for videocassette recorder), *CD* (for compact disk), *EC* (for European Community).

2.8.4 Acronyms

Acronyms are specially contrived to be used as spoken forms or read as words and clearly demonstrate their dependence on the spelling patterns of writing, e.g. *AIDS* (for acquired immune deficiency syndrome), OPEC (for Organization of Petroleum Exporting Countries, with omission of *of*).

2.8.5 Substitution

Another process of abbreviation rooted in written forms is substitution, which often involves abbreviated forms of foreign words and expressions that are adopted in English contexts. Latin abbreviations are included in substitution, such as *tid, qid, cf, e.g., i.e.*

2.7.6 Contraction

Contraction is a shortened form of a word or words. The mechanisms of contractions are largely rooted in written English and they produced only occasional forms in spoken English, e.g. *Dr.* (for Doctor).

2.9 Lexicography and Technical Dictionaries

2.9.1 Lexicography

The term lexicography means compiling dictionaries and the dictionary is defined as reference book or list of words (usually in alphabetical order) together with a guide to their meanings, pronunciation, spelling, or equivalents in other languages. The purpose of lexicography is the production of dictionaries, and dictionaries deal among other things with the ever-changing meanings of words. Lexicography is thus concerned with the job of describing all or some of the words of one or more languages in terms of their characteristic features, notably of their meaning.

There are two types of lexicography, namely *general lexicography* and *specialized lexicography. General lexicography* focuses on the design, compilation, use and evaluation of general dictionaries, i.e. dictionaries that provide a description of the language in general use. Such a dictionary is usually called a general dictionary or English for general purposes. *Specialized lexicography* is a part of general lexicographical theory and practice (Bergenholtz and Tarp 1995: 11). In recent decades, the further development of lexicographical theory is primarily due to linguists, with experts within the individual subject areas participating in the actual compilation of technical dictionaries. Specialized lexicography focuses on the design, compilation, use and evaluation of specialized dictionaries, i.e. dictionaries that are devoted to a (relatively restricted) set of linguistic and factual elements of one or more specialist subject fields, e.g. legal lexicography. Such a dictionary is usually called a specialized dictionary or technical dictionary. In short, a specialized

dictionary is a dictionary that covers a relatively restricted set of phenomena and often referred to as a technical dictionary.

2.9.2 Technical dictionaries

Technical dictionaries are devoted to the description of the technical language of a specialized subject or discipline. They are compiled on the basis of criteria provided by particular target groups and their professional or special-interest needs. They are deal with the language of a technical area such as engineering or a trade. The technical dictionary may be professional in serving legal, medical, clerical and scientific personnel in their work and it may also be aimed at specific academic fields of study: musicology, geology, philosophy, or even literature (Opitz 1983: 163).

The demand for all kinds of technical dictionaries goes hand in hand with the development of technology. Any mention of the technical dictionary will rise in the minds of most people associations of modern sophisticated engineering and industrial processes. This need continues in the twentieth century. Fields as diverse as medical, law, or engineering have been covered by technical dictionaries. It is governed by some specificity of purpose and represents a variety based on a classification of a subject matter.

2.9.2.1 The Ordinary Technical Dictionaries

There are many types of ordinary technical dictionaries designed to cover one segment of the whole lexicon of language. They diverge also on whether they are monolingual, bilingual or multilingual. Like their general dictionary, the greatest number of bilingual technical dictionaries published today have been designed to provide information primarily in the form of lexical equivalents or synonyms, with a view to translation into and from the native language. The headword of the entry generally consists of a single word and multi word variants of the term. It provides definitions as well as synonyms. It is assumed that the users are consulting the dictionary to find the equivalent in their mother tongue. Unfortunately, most of the existing ordinary technical dictionaries focus on providing definitions, but ignore other information useful for EFL students, for instance, the grammatical category, the phonetic description, examples usage, and illustrations (Opitz 1983: 106).

2.9.2.2 The Corpus-based Technical Dictionaries

Oxford University Press was the first to develop *Technical Learners' dictionaries* of 1990s. The first dictionary was *The Oxford Dictionary of Business English for Learners of English* (1993) edited by Allene Tuck. The intended learners are intermediate to advanced level who need to use English for both decoding and encoding in business. It was compiled with the combined expertise of EFL lexicographers and business. Firstly, it provides a range of linguistic information apart from definition, which existing monolingual technical dictionaries do only very infrequently. The grammar of words is explained, along with pronunciation and authentic examples showing how words are really used. Secondly, the dictionary was compiled with the help of the British National Corpus (BNC) a collection of 100,000,000 words of both written and spoken English using sub-disciplines of business for technical dictionaries. The publication of the *'The Oxford Dictionary of Business English for Learners of English'* represented linguistic advances in technical dictionary-making. In terms of the other features, the dictionary provides clear and useful information so that the EFL/ESP students can use the terms successfully for both decoding and encoding activities.

The above example should be taken into account in the preparation of technical dictionaries in the future. The compilation of the corpus should be determined by the precise purpose for which it intended to be used. Then, the lexicographer can compile technical dictionaries to satisfy the lexical demands on learners and consult with the expert for technical meanings.

In short, technical dictionaries should be based on a detailed analysis of how the terms are really used, with the aid of computer facilities to examine the words in authentic texts. Technical dictionaries should include information on word class, syntax and word formation. They should also provide pronunciation either in the form of phonetic transcription or by indication of stress patterns, the inclusion of examples also ensures that the user is given a certain amount of linguistic and encyclopedic information. Indeed we need to have illustrations in ESP dictionaries to make them more attractive and informative.

CHAPTER 3

METHODOLOGY

Research methodology in this chapter is divided into five sections. The first section includes the framework for the analysis. In the second section, data collection is presented. Research tools are described in the third section. The fourth section explains how to analyze the data. In the final section, the design of the sample medical dictionary is provided.

3.1 Framework for the Analysis

In order to reach the goals and objectives of the study, the following stages were carried out.

1. The pharmaceutical leaflets of drugs related to the first five causes of illness and death were selected as the primary input for the pharmaceutical corpus.
2. The contents of the pharmaceutical leaflets were transcribed into the computer readable forms by using the scanning process.
3. The concordancing software 'Wordsmith Tools' was utilized to create word frequency lists, count the number of tokens and types, and calculate the types/tokens ratio.
4. The word frequency lists were divided into three groups, namely general vocabulary, academic vocabulary, and technical vocabulary.
5. The three groups of vocabulary were classified into nouns, verbs, adjectives, and adverbs.
6. The abbreviations and collocations were identified.
7. The sample medical dictionary entries were designed.
8. The headwords of sample dictionary entries were selected from the ten most frequent words from the list of technical vocabulary, abbreviations and collocations.

3.2 Data Collection

Three hundred pharmaceutical leaflets were collected as a sample in this study. The statistic of the first five causes of illness and death in Thailand from the Public Health statistics (Ministry of Public Health 2000) was used as guidelines for the data collection in this study. Of the 300 leaflets, 150 leaflets were drawn on the basis of the first five causes of illness in Thailand: diseases of respiratory system, diseases of digestive system, diseases of musculoskeletal system and connective tissue, disease of skin and subcutaneous tissue system, and disease of circulatory system. Another 150 leaflets were collected on the basis of the first five causes of death in Thailand: malignant neoplasm and all forms; diseases of circulatory system; pneumonia and other diseases of lung; nephritis, nephrotic syndrome and nephrosis; and disease of liver and pancreases. As the circulatory system disease is the cause of both illness and death, the 300 leaflets were regrouped into nine categories derived from the causes of illness or death mentioned above, as shown in Table 3.1.

Table 3.1 The Number of the Collected Pharmaceutical Leaflets

Categories	Number of leaflets	Percent (%)
1. Disease of circulatory system	64	21.33
2. Disease of respiratory system	63	21.00
3. Malignant neoplasm, and all forms	62	20.67
4. Diseases of digestive system	34	11.33
5. Disease of musculoskeletal system and connective tissue	22	7.33
6. Disease of skin and subcutaneous tissue system	16	5.33
7. Nephritis, nephrotic syndrome and nephrosis	14	4.67
8. Pneumonia and other diseases of lung	14	4.67
9. Disease of liver and pancreases	11	3.67
Total	**300**	**100%**

Regarding a preliminary check on the pharmaceutical corpus size from the 300 leaflets, the corpus size is 344,914 running words which are large enough to be the corpus for the data analysis.

As claimed by Tribble and Jones (1990), the number of running words should be at least 50,000 words for a specialized corpora.

3.3 Research Tools

3.3.1 Concordancing Software

Concordancing software is a primary tool in all corpus study, and enables a researcher to find all occurrence of a given word, a part of word, or combinations of words within the context in a corpus of texts (Simon 1996: 190).

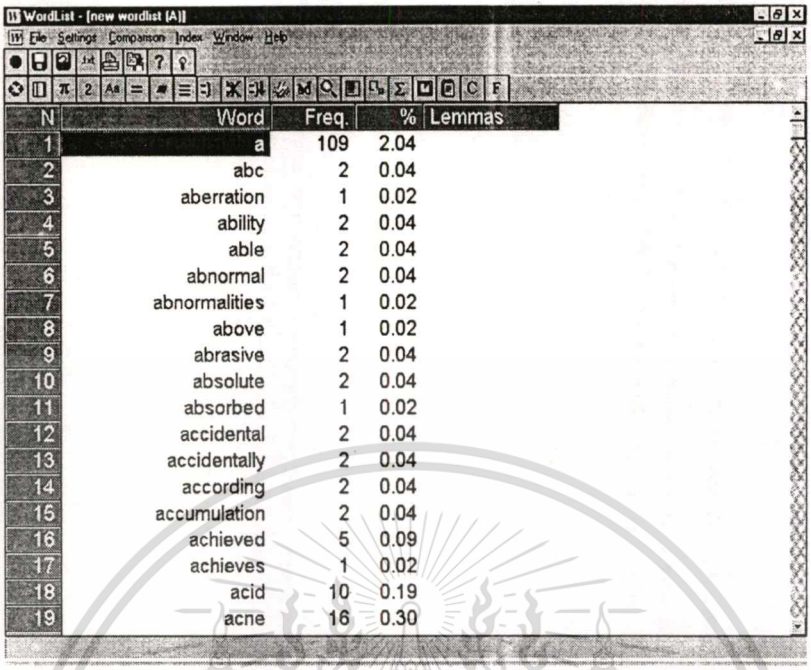
According to Boswood (1997: 225), a concordancer should be easy to use and should be able to (1) produce numerically and alphabetically sorted word lists; (2) produce concordances in which the contexts can be sorted to the right or the left of the key word; (3) handle significant quantities of text: (4) work quickly; and (5) produce results directly to the screen where reviewing and editing can be carried out.

There are many well known concordancing programs. Examples include:

- *Longman Mini Concordancer* is a program which works with words on plain ASCII text.
- *Microconcord* is a concordancer for the PC and appropriate for pedagogic applications.
- *Wconcord* is a freeware concordancing program for Windows.
- *WordCruncher* is a program that can produce frequency lists of corpora and key word in context displays, word combinations and parts of words.
- *WordSmith Tools* assists in the text analysis of either a single or a large text.

In this study, the concordancing software called *WordSmith Tools Version 3.0* was used for the corpus analysis. This program was created by Mike Scott in 1996 and has been distributed by Oxford University Press. The Wordsmith Tools consist of three main tools, namely Concord Tool, Keywords Tool, and Wordlist Tools. The Wordlist Tool and Concord Tools were used to assist the vocabulary analysis in this study.

The Wordlist Tool is used to generate word lists. The word frequency lists were automatically presented in either alphabetical order or frequency order, as shown in Figures 3.1 and 3.2.

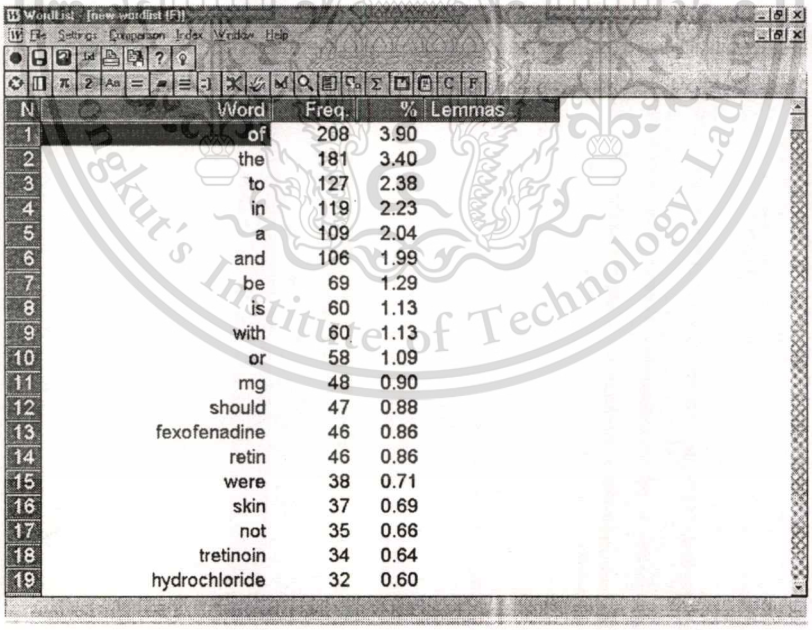


WordList - [new wordlist (A)]

File Settings Comparison Index Window Help

N	Word	Freq.	%	Lemmas
1	a	109	2.04	
2	abc	2	0.04	
3	aberration	1	0.02	
4	ability	2	0.04	
5	able	2	0.04	
6	abnormal	2	0.04	
7	abnormalities	1	0.02	
8	above	1	0.02	
9	abrasive	2	0.04	
10	absolute	2	0.04	
11	absorbed	1	0.02	
12	accidental	2	0.04	
13	accidentally	2	0.04	
14	according	2	0.04	
15	accumulation	2	0.04	
16	achieved	5	0.09	
17	achieves	1	0.02	
18	acid	10	0.19	
19	acne	16	0.30	

Figure 3.1 Alphabetical Lists



WordList [new wordlist (F)]

File Settings Comparison Index Window Help

N	Word	Freq.	%	Lemmas
1	of	208	3.90	
2	the	181	3.40	
3	to	127	2.38	
4	in	119	2.23	
5	a	109	2.04	
6	and	106	1.99	
7	be	69	1.29	
8	is	60	1.13	
9	with	60	1.13	
10	or	58	1.09	
11	mg	48	0.90	
12	should	47	0.88	
13	fexofenadine	46	0.86	
14	retin	46	0.86	
15	were	38	0.71	
16	skin	37	0.69	
17	not	35	0.66	
18	tretinoin	34	0.64	
19	hydrochloride	32	0.60	

Figure 3.2 Frequency Lists

The Concord is used to make a concordance data. The users will specify a search word, which the Concord will seek in all the text files users have chosen. The program will then present a concordance data, and allow the researchers to access information about collocates of the search word, as shown in Figure 3.3.

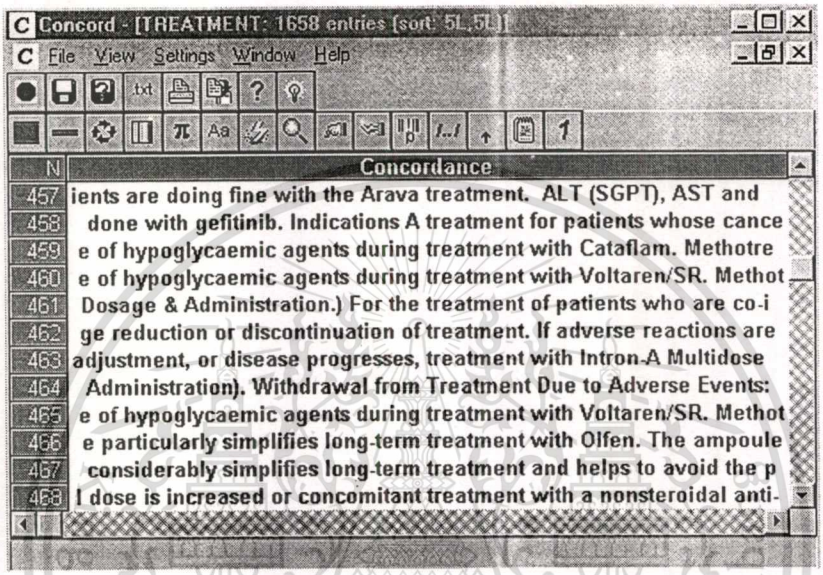


Figure 3.3 A KWIC Concordance: Entry of 'treatment'

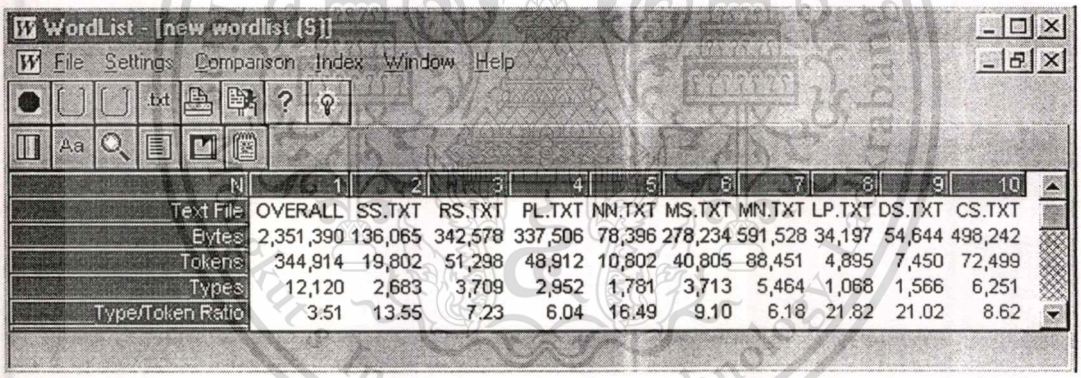
Figure 3.3 shows the concordance line of *'treatment'*. In Concord tool, the Keywords is used to locate and identify key words in a given text. To do so, it compares the words in the text with a reference set of words usually taken from a large corpus of text. Any word which is found to be outstanding in its frequency in the text is considered "key". It will then present a concordance display and information about collocates of the such word. The researchers can be able to see lots of examples of word or phrase in their context.

3.3.2 Medical Dictionary

The guidelines on the technical vocabulary analysis were mainly based on *Mosby's Medical, Nursing, & Allied Health Dictionary* (Anderson 2002). It was developed to provide a single source of authoritative, up-to-date information concerning health and health care terminology. Moreover, *Merriam Webster: Medical Dictionary Online* was also used to define the technical vocabulary. These dictionaries were chosen as tools for word selection and identification of technical vocabulary.

3.4 Data Analysis

All printed pharmaceutical leaflets were scanned and stored as plain-text files (*.txt). There were 9 text files in the pharmaceutical corpus, namely disease of skin and subcutaneous tissue system (SS.txt.), diseases of respiratory system (RS.txt), pneumonia and other diseases of lung (PL.txt), nephritis, nephrotic syndrome and nephrosis (NN.txt), diseases of musculoskeletal system and connective tissue (MS.txt), malignant neoplasm and all forms (MN.txt), disease of liver and pancreases (LP.txt), diseases of digestive system (DS.txt), and disease of circulatory system (CS.txt). At the scanning stage, the optical character recognition (OCR) software was used to recognize scanned letters and represent the text electronically. Each text file was exported into Microsoft Word 2000 as a word document (*.doc) and the spelling of all words in the text was checked by using the spell-checker. The Wordlist Tool was used to calculate the frequency of occurrence of running words or tokens and word types, including type/token ratio, as presented in Figure 3.4.


WordList - [new wordlist (5)]

File Settings Comparison Index Window Help

N	1	2	3	4	5	6	7	8	9	10
Text File	OVERALL	SS.TXT	RS.TXT	PL.TXT	NN.TXT	MS.TXT	MN.TXT	LP.TXT	DS.TXT	CS.TXT
Bytes	2,351,390	136,065	342,578	337,506	78,396	278,234	591,528	34,197	54,644	498,242
Tokens	344,914	19,802	51,298	48,912	10,802	40,805	88,451	4,895	7,450	72,499
Types	12,120	2,683	3,709	2,952	1,781	3,713	5,464	1,068	1,566	6,251
Type/Token Ratio	3.51	13.55	7.23	6.04	16.49	9.10	6.18	21.82	21.02	8.62

Figure 3.4 Statistical Analysis Details of the Pharmaceutical Corpus

The word lists were manually grouped into general, academic, and technical vocabulary. Each was categorized into nouns, verbs, adjectives, and adverbs. The abbreviations and collocations were identified.

3.5 The Design of the Sample Dictionary

This section is devoted to the rationale for the design of he sample dictionary. Zgusta (1971) identified the four main tasks of the dictionary maker as (1) the collection of material, (2) the selection of entries, (3) the construction of entries, and (4) the arrangement of the entries. For the

sample dictionary, the technical terms and the selection of headwords are collected from the pharmaceutical corpus and prioritized on grounds of overall frequency. The selection of the construction and internal arrangement of the entries are based on Landau (2001) suggested that a typical dictionary provides its users with various kinds of information about words: their meaning, pronunciation, etymology, part of speech and status. In addition, dictionaries will contain a series of example sentence to illustrate in a meaningful context to show the various meaning that a given word has. The sample dictionary entries should consist of headword, pronunciation, grammatical information, word root, English definition, Thai synonym, example of usage and illustration.

3.5.1 An Entry Structure of the Sample Dictionary

The entry provided below has a number of different fields, namely headword, pronunciation, grammatical information, English definition, Thai synonym, root, example of usage, and illustration.

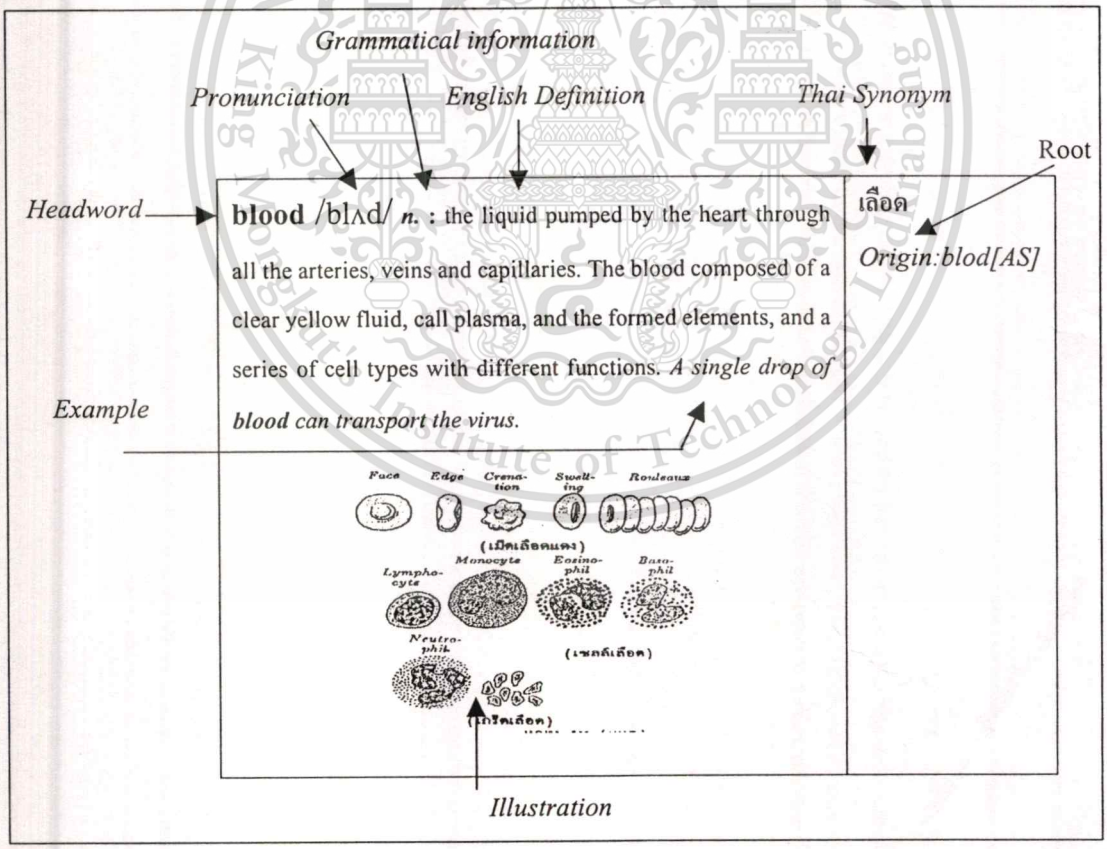


Figure 3.5 The Design of Sample Medical Dictionary Entry

3.5.1.1 Headwords

The headwords list for the sample dictionary starts with the content words or technical terms with the high frequencies in the pharmaceutical corpus. Only medical terms and abbreviations in corpus with the high frequency are considered for inclusion.

3.5.1.2 Usage in Context

Words or Terms in Context are a recognized feature of English language teaching. They provide information on collocations and grammatical information in authentic text. The sample dictionary must therefore provide a number of examples to show typical contexts that the word is likely to be used in. These examples are drawn from the concordance of the pharmaceutical corpus.

3.5.1.3 Grammatical Information

It is important that the technical dictionary should provide grammatical information. The grammatical information is more essential for the person who trying to speak or understand a foreign language than for the native speaker. So the sample dictionary provides grammatical information in terms of part of speech. A set of grammatical information in abbreviated forms was used, i.e. *n.* for noun, *v.* for verb, *adj.* for adjective, *adv.* for adverb.

3.5.1.4 Definitions

The sample dictionary gives Thai synonyms for the headword as well as the English definitions. The bilingual elements help to encourage students to read terms in contexts, and to develop resources in English. In this dictionary, English definitions provide by using the data from *Mosby's Medical, Nursing, & Allied Health Dictionary* and *Merriam Webster: Medical Dictionary Online.* Example of usage carried out from the data of corpus evidence. Thai synonyms are based on the Thai Royal Terminology (1996).

3.5.1.5 Pronunciation and Word Stress

Two kinds of information are needed to pronounce a word properly: the sounds and stress. In the sample dictionary, the pronunciation is based on the International Phonetic Alphabet (IPA) which is used virtually in standard monolingual dictionaries because IPA symbols are phonetically based on, i.e. the area and manner of articulation (Landau 2001: 128).

3.5.1.6 Etymology

Etymologies are often included in scientific and technical dictionaries, especially those for the older science such as medicine, because they are essential to an understanding of meaning (Landau 2001: 131). The sample dictionary entries provide the root of word, e.g. ***Origin**: blod [AS].*

3.5.1.7 Illustrations

Illustrations such as diagram, charts are used where their inclusions will be of practical value. They will enhance the student's understanding of the terms. Therefore, the sample dictionary provides illustrations and their labels in both Thai and English.

CHAPTER 4

RESULTS AND DISCUSSION

This chapter focuses on the results and discussion of vocabulary analysis used in the pharmaceutical leaflets, namely the statistical details of the whole corpus, the word frequency lists, vocabulary in the pharmaceutical corpus, abbreviations and collocations. The sample entries of the sample dictionary are included in this chapter.

4.1 Statistical Details of the Whole Corpus

The statistical details of the pharmaceutical corpus show a summary of the content of a corpus. Table 4. 1 summarizes the number of the tokens (running words), word types and types/token ratio. Type/token ratio is computed by dividing the number of tokens by the number of word types. The lower type/token ratio is, the greater the diversity of words in the corpus is.

Table 4.1 The Statistical Details of the Whole corpus

Tokens	344,914
Word types	12,057
Type/token Ratio	1: 28.61

Table 4.1 shows the corpus size consisting of 344,914 tokens or running words. The total number of word types equals 12,057 words in the corpus since a recurrent word is counted only once. The ratio of types/tokens was 1: 28.61. The ratio indicated each word is repeated nearly 29 times on average throughout the corpus.

4.2 The Word Frequency Lists

This section shows three main word frequency lists. First, it shows the twenty most frequent words of the whole corpus (see Appendix C). Second, the word frequency list of the two specialized corpora, i.e. pharmaceutical corpus and biology corpus, are compared with the other three general

corpora, i.e. Brown corpus, LOB Corpus, and COUBILD Corpus. Third, it shows the word frequency list of the 10 high-frequency content words from the pharmaceutical corpus.

Table 4.2 The Twenty Most Frequent Words and Abbreviaitions in the Pharmaceutical Corpus

Rank	Word	Frequency	% of occurrence
1.	the	12,699	3.68
2.	of	12,530	3.63
3.	in	9,167	2.66
4.	and	8,572	2.49
5.	to	6,016	1.74
6.	with	5,053	1.47
7.	be	4,511	1.31
8.	is	4,259	1.23
9.	or	4,116	1.19
10.	a	4,055	1.18
11.	patients	3,250	0.94
12.	mg	2,685	0.78
13.	should	2,616	0.76
14.	for	2,524	0.73
15.	may	2,059	0.60
16.	dose	1,894	0.55
17.	not	1,860	0.54
18.	treatment	1,656	0.48
19.	as	1,642	0.48
20.	been	1,581	0.46

Table 4.2 shows the twenty most frequent words and abbreviations in the pharmaceutical corpus. The function words occur in the top ten rank. The functions words are *'of', 'in', 'and', 'to',*

'with', *'be'*, *'is'*, *'or'*, and *'a'*. Within the pharmaceutical corpus, the ten most frequently occurring types account for 20.58 % of the total list. The function word *'the'* occured the most frequently with 12,699 times or 3.68 %. The content word, *'patients'* appeared the highest with 3,250 occurrences or 0.94 %. The content words from the Table 4.2, namely *'patient'*, *'dose'*, and *'treatment'* are highly medicine-related to the subject matter of the corpus. The first abbreviation *'mg'* occurs in the rank 12nd.

The following table shows the comparison of general and specialized corpora. The Brown, LOB, and COUBILD are general corpora, whereas the biology and pharmaceutical corpus are specialized corpora.

Table 4.3 The Ten Most Frequent Words in the General Corpora and the Specialized Corpora

Rank	General Corpora			Specialized Corpora	
	Brown	*LOB*	*COUBILD*	*Biology*	*Pharmaceutical*
1.	the	the	the	the	the
2.	of	of	of	and	of
3.	and	and	to	of	in
4.	to	to	and	is	and
5.	a	a	a	a	to
6.	in	in	in	in	with
7.	that	that	that	are	be
8.	is	is	's	to	is
9.	was	was	is	it	or
10.	he	it	it	this	a

Table 4.3 shows the comparison of three general corpora and two specialized corpora. The ten most frequent word are function words, namely articles, prepositions, the verbs to be, conjunctions, and determiners, for instance *'the'*, *'of'*, *'is'*, *'and'*, *'in'*. These results are in substantial agreement with the study of Flowerdew (1993) that both general corpora and specialized corpora share similar aspect in that most 10 frequent words. The function words are used in the texts of

different disciplines. The word with highest frequency in the general and specialized corpora is the word *'the'*.

The next table illustrates the 10 high-frequency content words in the pharmaceuctical corpus.

Table 4.4 The 10 high-frequecy content words and abbreviations in the Pharmaceutical Corpus

No.	Rank	Words	Freq.	%
1.	11	patients	3,250	0.94
2.	12	mg	2,685	0.78
3.	16	dose	1,894	0.55
4.	18	treatment	1,656	0.48
5.	26	therapy	1,160	0.34
6.	27	administration	1,076	0.31
7.	28	use	1,061	031
8.	29	dosage	1,037	0.30
9.	36	ml	936	0.27
10.	37	daily	924	0.27

With regard to the content words and abbreviations in the pharmaceutical corpus, Table 4.4 shows the ten most frequent content words, namely *'patients'*, *'mg'*, *'dose'*, *'treatment'*, *'therapy'*, *'administration'*, *'use'*, *'dosage'*, *'ml'*, and *'daily'*. We can see that most of these words, e.g *'dose'*, *'therapy'*, *'dosage'* function as nouns which give information about the use of drug or medication. In short, the findings of this analysis reflect the lexical feature of the pharmaceutical corpus.

4.3 Vocabulary in Pharmaceutical Corpus

The purpose of the study is to analyze the vocabulary in the pharmaceutical corpus by categorizing them into three main groups, namely general, academic, and technical vocabulary. From the total number of 12,057 word types, they were analyzed in terms of a word frequency list. The vocabulary in the word frequency list was drawn to frequency 35 (0.01%) of occurrence. Then, the

word frequency list was lemmatized. Eventually, the word frequency list consists of 276,057 tokens or 1,223 lemmas. The 1,223 lemmas were categorized into 4 groups, i.e. general, academic, technical vocabulary and other kinds of words outside the three main types. The results were shown in Table 4.5 below:

Table 4.5 Types of Vocabulary in the Pharmaceutical Corpus

Types of Vocabulary	Lemmas	%
General	473	38.68
Academic	113	9.24
Technical	404	33.03
Others	233	19.05
Total	**1,223**	**100**

Table 4.5 indicates that the four types of vocabulary in the pharmaceutical corpus. 473 lemmas or 38.68 % are general vocabulary, 113 lemmas or 9.24 % are academic vocabulary, and 404 lemmas or 33.03 % are technical vocabulary, and 233 lemmas or 19.05 % are the other kinds of words outside the three main types, i.e. trademarks or brand names, distributors, and abbreviations. The examples of names of trademarks in the pharmaceutical corpus are *'Ventolin'*, *'Herceptin'*, *'Xloda'*, *'Nizoral'*, and etc. The names of distributors in the pharmaceutical corpus are *'Zueling'*, *'Diethelm'*, and so on. The abbreviations in the pharmaceutical corpus are *'tid'*, *'MAOIs'* (See section 4.4).

The percentage of the types of vocabulary in the pharmaceutical corpus was compared with those of the GSL and the AWL. Nation (2000) estimated that in the general standard text, there are 80% of the GSL, 10% of the AWL, and 10 % of the words outside the GSL and the AWL (5% of technical terms and 5% of the words outside those three groups). In the pharmaceutical corpus, 404 lemmas or 33.03% are technical vocabulary. This showed that language of medical leaflets is highly technical by nature. This finding confirmed by the work of Method (1997: vi) that medical wirting relied very heavily on a technical vocabulary.

4.3.1 General Vocabulary in the Pharmaceutical Corpus

Of the 1,223 lemmas, the general words are consisted of 473 words or 38.68 %. Of the 473 general words, they consist of 105 function words and 368 content words (see Appendix D for more details). Table 4.6 shows the top twenty general vocabulary in the pharmaceutical corpus. They are grouped into 4 main parts of speech: nouns, verbs, adjectives, and adverbs.

Table 4.6 The Top Twenty General Vocabulary in the Pharmaceutical Corpus

Nouns		Verbs		adjectives		adverbs	
1. Patient	3,583	1. Use	1,456	1. Daily	942	1. Usually	288
2. Treatment	1,677	2. Increase	1,096	2. Severe	636	2. Rarely	243
3. Effect	1,615	3. Report	804	3. High	517	3. Elderly	241
4. Day	1,108	4. See	763	4. Low	430	4. Well	228
5. Hour	923	5. Recommend	660	5. Following	373	5. Generally	156
6. Blood	789	6. Include	637	6. Single	255	6. Respectively	142
7. Study	788	7. Give	566	7. Great	255	7. Immediately	121
8. Time	672	8. Treat	557	8. Long	198	8. Especially	117
9. Child	560	9. Observe	515	9. Active	198	9. Carefully	93
10. Level	553	10. Take	465	10. Possible	191	10. Rapidly	81
11. Case	525	11. Receive	453	11. Rare	193	11. Mainly	76
12. Week	488	12. Control	421	12. Mild	191	12. Weekly	74
13. Use	446	13. Reduce	416	13. Small	186	13. Closely	64
14. Action	445	14. Contain	382	14. Effective	160	14. Slightly	60
15. Year	438	15. Cause	379	15. Least	152	15. Completely	45
16. Pack	426	16. Decrease	353	16. Total	152	16. Highly	44
17. Solution	419	17. Show	349	17. Moderate	149	17. Slowly	43
18. Pressure	405	18. Following	310	18. Serious	141	18. Possibly	40
19. Agent	387	19. Associate	289	19. Common	138	19. Commonly	40
20. Result	368	20. Discontinue	280	20. White	137	20. Clearly	37

As can be seen in Table 4.6, the high frequency of nouns in the phamaceutical corpus are *'patient'* (freq. 3,583), *'treatment'* (freq. 1,677), 'effect' (freq. 1,615). Their meanings are related to medical area. The findings of many nouns in the pharmaceutical corpus showed that the pharmaceutical language is concerned with the high density of information and infomative by nature.

The general verbs such as ***'use', 'increase', 'report', 'recommend'*** appear in the past form such *as 'reported', 'recommended', 'increased'* and so on. For *example, 'Deaths have been* ***reported****', 'Cough has been* ***reported*** *with the use of ACE inhibitors'*. Since the message in the pharmaceutical leaflets are always written in the form of report, caution, and suggestion using the passive voice (to be+V_3), as shown in the following concordance data below:

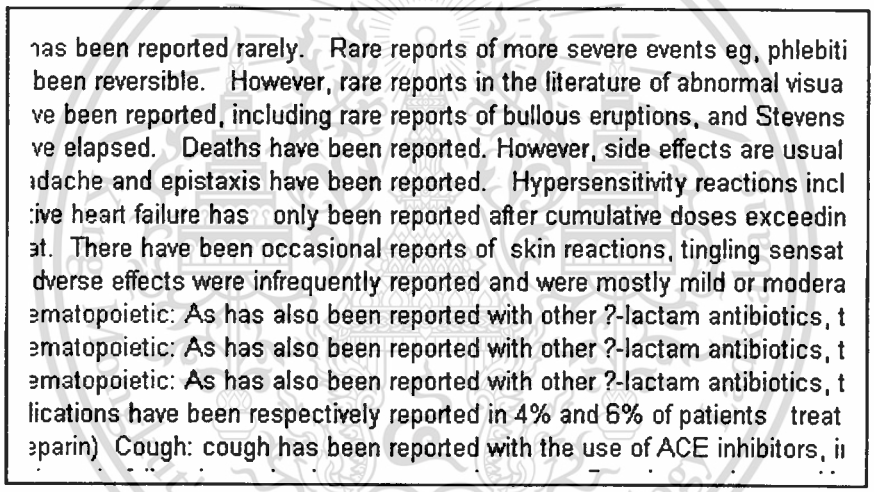


Figure 4.1 The Concordance Data of the Word 'report'

The overlap between nouns and verbs can be identified by using the concordance data from the corpus. For example, the word ***'use* (1,902)'** (see Appendix D) has more than one function which can be nouns or verbs. The concordance data in Figure 4.2 show clearly most of the words ***'use'*** are verbs. There are 1,456 of the word 'use' belonging to verbs, e.g. *'Patients of reproductive age should* ***use*** *contraceptives during therapy and for the following 6 mths'*. The other 446 belong to nouns, e.g. *'The* ***use*** *of this medicine is contraindicated during the second and third trimesters of pregnancy'*.

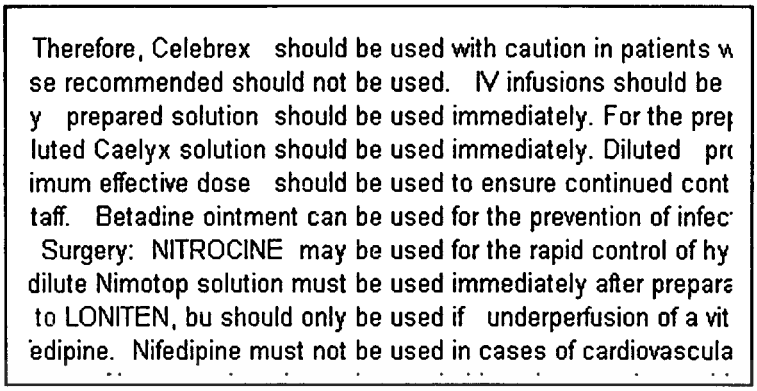


Figure 4.2 The Concordance Data of the Verb 'use'

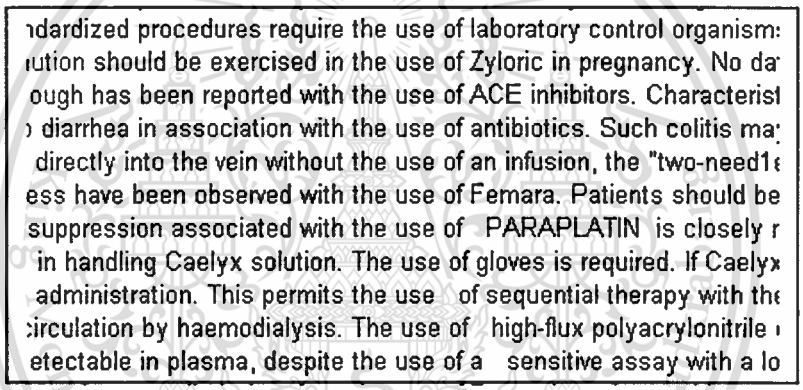


Figure 4.3 The Concordance Data of the Noun 'use'

The overlap between adjectives and verbs can be identified by using concordance data from the corpus. For example, the word ***'following* (683)'** (see Appendix D) has more than one function which can be adjectives or verbs. In this study, we found 373 of adjectives and 310 of verbs for the word *'following'*.

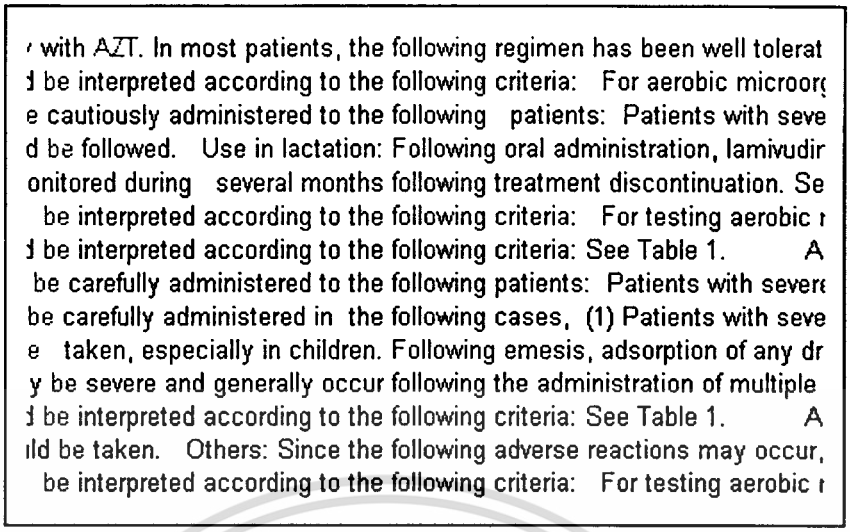


Figure 4.4 The Concordance Data of the Adjective 'following'

The concordance data in Figure 4.4 show clearly most of the words ***'following'*** are adjectives for example, *'The **following** therapeutic measures should be taken in cases of overdosage', 'In clinical trials, the **following** events occurred in 1-3% of patients, regardless of drug relationship'*. An example of verb 'following' is *'It is **following** a parenteral dose of equal size'*.

4.3.2 Academic Vocabulary in the Pharmaceutical Corpus

Of the 1223 lemmas, there are 113 Lemmas or 9.24 % of academic words. Of the 113 academic words, they consist of 1 function word, i.e. *'via'* (freq. 80) and 112 content words (See Appendix E). The content words are identified into nouns, verbs, adjectives, and adverbs. Table 4.7 shows the top twenty academic vocabulary in the pharmaceutical corpus.

Table 4.7 The Top Twenty Academic Vocabulary in the Pharmaceutical Corpus

Nouns		Verbs		Adjectives		Adverbs	
1. Administration	1,083	1. Occur	864	1. Normal	250	1. Approximately	305
2. Concentration	622	2. Require	396	2. Potential	245	2. Significantly	101
3. Function	524	3. Indicate	306	3. Significant	243	3. Primarily	61

Table 4.7 (continued)

Nouns		Verbs		Adjectives		Adverbs	
4. Adult	242	4. Range	267	4. Appropriate	197	4. Initially	56
5. Monitor	368	5. Inhabit	201	5. Maximum	189	5. Potentially	54
6. Indication	357	6. Establish	181	6. Prior	155	/	
7. Data	231	7. Induce	179	7. Initial	149	/	
8. Incidence	185	8. Affect	131	8. Similar	149	/	
9. Elimination	176	9. Initiate	122	9. Reversible	134	/	
10. Inhibitor	176	10. Obtain	109	10. Annual	128	/	
11. Phase	170	11. Adjust	103	11. Specific	116	/	
12. Evidence	164	12. Enhance	98	12. Individual	106	/	
13. Duration	157	13. Exceed	98	13. Medical	104	/	
14. Maintenance	134	14. Maintain	91	14. Adequate	103	/	
15. Physician	134	15. Evaluate	86	15. Subsequent	96	/	
16. Adjustment	133	16. Isolate	86	16. Major	75	/	
17. Interval	128	17. Involve	82	17. Primary	66	/	
18. Insufficiency	120	18. Alter	74	18. Concurrent	61	/	
19. Volume	118	19. Conducted	71	19. Apparent	56	/	
20. Site	117	20. Detect	58	20. Visual	51	/	

Table 4.7 shows the academic vocabulary with high frequency in the pharmaceutical corpus. The academic adverbs are rarely found. The words such as *'administration'*, *'occur'*, *'normal'*, *'approximately'* are found in the academic fields.

4.3.3 Technical Vocabulary in the Pharmaceutical Corpus

Of the 1223 lemmas, there are 404 words or 33.03 % are technical vocabulary. Of the 404 technical words, they are identified into nouns, verbs, adjectives, and adverbs (See Appendix F). The

technical vocabulary list consists of 293 nouns, 16 verbs, 90 adjectives, and 5 adverbs. Table 4.8 shows the top twenty technical vocabulary in the pharmaceutical corpus.

Table 4.8 The Top Twenty Technical Vocabulary in the Pharmaceutical Corpus

Nouns		Verbs		Adjectives		Adverbs	
1. Dose	2,866	1. Administer	720	1. Adverse	790	1. Clinically	167
2. Drug	1,358	2. Impair	270	2. Renal	779	2. In vitro	165
3. Therapy	1,170	3. Excrete	169	3. Clinical	689	3. Orally	137
4. Dosage	1,076	4. Inhale	127	4. Oral	576	4. Concomitantly	93
5. Tablet	986	5. Dilute	91	5. Acute	408	5. In vivo	59
6. Plasma	620	6. Propionate	88	6. Hepatic	395	/	
7. Symptom	530	7. Retard	81	7. Cardiac	284	/	
8. Pregnancy	501	8. Potentate	79	8. Chronic	250	/	
9. Infusion	464	9. Contraindicate	78	9. Gastrointestinal	235	/	
10. Precaution	462	10. Prescribe	75	10. Therapeutic	197	/	
11. Serum	393	11. Assay	66	11. Concomitant	191	/	
12. Liver	369	12. Blurr	52	12. Nasal	189	/	
13. Disorder	326	13. Metabolise	52	13. Allergic	184	/	
14. Corticosteriod	321	14. Discard	44	14. Pregnant	176	/	
15. Impairment	318	15. Precipitate	38	15. Urinary	162	/	
16. Contraindication	310	16. Activate	37	16. Peripheral	154	/	
17. Hypersensitivity	294	/		17. Inflammatory	149	/	
18. Cancer	286	/		18. Respiratory	143	/	
19. Toxicity	274	/		19. Multidose	125	/	
20. Metabolite	258	/		20. Infant	120	/	

Table 4.8 shows that the most frequent technical vocabularies are nouns and adjectives. Verbs and adverbs are rarely found in the corpus. Most medical terms have Greek or Latin origins, e.g. hypertension(prefix-hyper), dermatitis(root-derm(at)).

4.4 Abbreviations

There are 95 abbreviations in the whole corpus (See Appendix G). The top twenty abbreviations in the corpus are as follows:

Table 4.9 The Top Twenty Abbreviations in the Pharmaceutical Corpus

No.	Abbreviations	Freq.	Full Words
1.	mg	2,685	Milligram
2.	ml	936	Milliliter
3.	kg	549	Kilogram
4.	eg	513	For example (Latin : exempli gratia)
5.	mcg	484	Microgram
6.	tab	472	Tablet
7.	IV	453	Intravenous
8.	C	309	Celsius/centigrade
9.	min	306	Minute
10.	hr(s)	245	Hour/hours
11.	l	216	Liter
12.	g	189	Gram
13.	w/	170	With
14.	p	164	Pressure
15.	s	143	Staphylococcus
16.	MIMS	139	Monthly Index of Medical Specialties
17.	mr	136	Milliroentgen
18.	HCl	120	Hydrochloride
19.	cap(s)	113	Capsule(s)
20.	i	108	I

In the pharmaceutical corpus, it is found that there are five main types of abbreviations: clipping, initialism, acronym, contraction and substitution. The findings are shown below:

1. **Clipping**: **tab**(tablet), **cap**(capsule), **min**(minute),
2. **Initialism**: **mg**(milligram), **ml**(milliliter), **kg**(kilogram), **mcg**(microgram), g(gram), w/(with), C(Celsius/centigrade), l(Liter) , w/(with) ,**s**(staphylococcus), **p**(pressure), **HCl**(Hydrochloride), **IV**(intravenous)
3. **Acronym**: **MIMS** (Monthly Index of Medical Specialties)
4. **Substitution: e.g.**(for example), i(1)
5. **Contraction**: **hr(s)**(hour(s)), **childn**(children), **yr**(years)

If the above abbreviations are closely considered, they can be grouped into broad areas of general and technical abbreviations. Firstly, the general abbreviations are commonly used in every field, such as *'mg'* (milligram), *'tsp'* (teaspoonful), *'hr/ hrs'* (hour/hours), *'e.g.'* (for example). Figure 4.5 shows the concordance data of the general abbreviation *'e.g.'*:

re observed, proper measures eg, dose reduction or discontin
ing exposure to various stimuli eg, exercise, cold air and the e
ecreased bone mineral content eg, chronic alcohol and/or tob
nerative osteoarticular disease eg, arthrosis and osteoarthritis
:ood. Anti-inflammatory actions eg, inhibition of inflammatory
CNS and endocrine parameters (eg, GH, prolactin, LH, testost
/e rise to clinical complications eg, obstructive jaundice, ascen
en and energy-rich phosphates eg, ATP and creatinine phosp
: Severe (Grade 3 or 4) toxicity eg, pleural effusion, pericardial
actions: Red cell abnormalities (eg, haemolytic anaemia, macr

Figure 4.5 Concordances of the General Abbreviation for 'e.g.'

The abbreviation ***'e.g.'*** appears in many areas. It is the mark of inductive reasoning as the writer uses to support any generalization with the specific case.

Secondly, technical abbreviations have highly specific meanings which are really difficult to understand for person who does not know or study medicine area. So the abbreviations such as

'IV' (intravenous), *'HCl'* (hydrochloride), *'GI'* (Gastrointestinal), *'NSAIDs'* (non-steridal anti-inflammatory drugs) need to be added in the medical science dictionary. Figure 4.6 shows the concordance data of technical abbreviation *'IV'*:

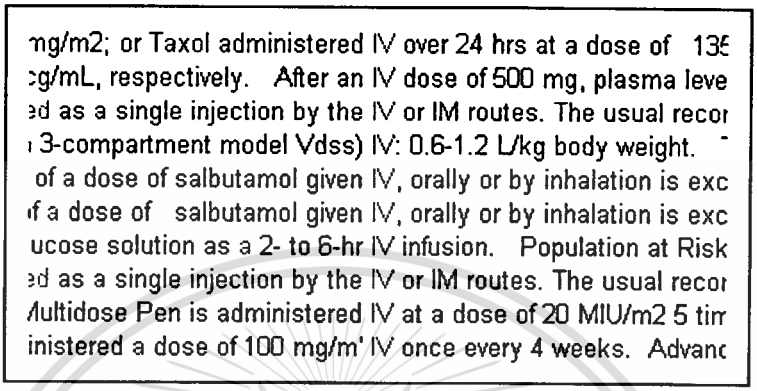


Figure 4.6 Concordances of the Technical Abbreviation for 'IV'

4.5 Collocations

As stated in Chapter 2, there are two types of collocations, i.e. lexical and grammatical collocations. They are phrases consisting of dominant words (nouns, adjectives, or verbs) and other elements. Table 4.10 shows the frequent lexical collocations in the pharmaceutical corpus. The following table shows the frequent Grammatical collocations in the pharmaceutical corpus.

Table 4.10 The Lexical Collocations in the Pharmaceutical Corpus

No.	Word	Lexical Collocations	Examples
1.	Dose	Daily dose	*The **daily dose** can be increased gradually to 60 mg.*
		Recommended dose	*The **recommended dose** is 2-4 mg once daily.*
		Single dose	*This may be taken as a **single dose** or as 2 divided doses.*

Table 4.10 (continued)

No.	Word	Lexical Collocations	Examples
2.	Treatment	Concomitant treatment	*Concomitant treatment may sometimes be necessary to control eye symptoms caused by the allergy.*
		Symptomatic treatment	*Observation should be carried out with symptomatic treatment if necessary.*
		Palliative treatment	*This cytotoxic action is the basis for their use in the palliative treatment of certain types of cancer.*
3.	Therapy	Corticosteroid therapy	*Corticosteroid therapy may alter the motility and number of spermatozoa.*
		Maintenance therapy	*For some patients, maintenance therapy may be achieved with less frequent application.*
		Combination therapy	*Renal function must be measured during combination therapy.*
4.	Drug	Drug interaction(s)	*No formal drug interaction studies have been conducted with Caelyx.*
		Unchanged drug	*An administered dose was recovered as unchanged drug in urine within 48 hrs.*
5.	Adverse	Adverse drug reaction	*Adverse drug reactions are usually mild.*
		Adverse event(s)	*The adverse event profile for patients who received Taxol in combination.*
		Adverse effects	*There were no adverse effects on behaviour or reproductive performance.*

Table 4.10 (continued)

6.	Renal	Renal function	*Monitoring of **renal function** is recommended as a precautionary measure.*
		Renal impairment	*Patients with **renal impairment** should not require adjustment of dosage.*
		Renal failure	*Patients with **renal failure** should be monitored while on Mucoflux.*
7.	Clinical	Clinical trial(s)	*All 10 women in the **clinical trial** reported breakthrough bleeding.*
		Clinical studies	*In **clinical studies** of patients with chronic hepatitis B, Zeffix was well tolerated.*
		Controlled clinical studies	***Controlled clinical studies** have shown that intranasal corticosteroids may cause a reduction.*
8.	Plasma	Peak plasma concentration(s)	***Peak plasma concentration** are reduced by approximately 98% within 3-4 hrs.*
		Plasma level(s)	***Plasma levels** of anticonvulsants may become subtherapeutic during cisplatin therapy.*
		Plasma clearance	***Plasma clearance** is on average 8 ml/min.*

Table 4.1 The Grammatical Collocations in the Pharmaceutical Corpus

No.	Word	Grammatical Collocations	Examples
1.	in	in patient with	*Trental 400 must not be employed **in patients with** hypersensitivity to pentoxifylline.*
2.	on	effect on	*Food had no significant **effect on** either the peak plasma concentration.*
3.	of	treatment of	*Symptomatic **treatment of** vertigo due to vestibular disorders.*

Table 4.11 (continued)

No.	Word	Grammatical Collocations	Examples
4.	with	associated with	*Conditions **associated with** prolonged circulation time.*
5.	for	recommended for	*Durogesic is not **recommended for** use in nursing women.*
6.	into	divided into	*In general, the daily dose is **divided into** 2-3 individual administrations.*
7.	about	about + (number)	*Binding of topotecan to plasma proteins **is about** 35%.*

The examples of lexical and grammatical collocations are taken from the concordance data of the pharmaceutical corpus. The following example in Figure 4.7 shows the lexical collocations.

ıerefore varies as a function of the daily dose. Complete remission of t
is usually 150 mg. In general, the daily dose is divided into 2-3 individ
steroid is started by reducing the daily dose gradually. Transferred pati
3. Generally speaking, the daily dose is 3 caps of Adalat or Ada
Occasionally, an increase of the daily dose to 900 mg may be necess
3 or 4 equally spaced doses- The daily dose of 36 mg should not be e:
ı improvement after 1 month, the daily dose may be gradually increas
ვ. In well-controlled patients, the daily dose may be reduced to <400 r
herapy, it is advisable to raise the daily dose according to the following
ənt diabetes mellitus patients, the daily dose is 10 mg 'Zestril' once da
. In severe cases (eg, colic), the daily dose can exceptionally be incre
m. In such cases, doubling of the daily dose may be necessary. The
ıpaired liver function as well, the daily dose must therefore be individu
ally. If the effect is insufficient, the daily dose may then be increased gr

Figure 4.7 Concordance Data for Lexical Collocations 'daily dose'

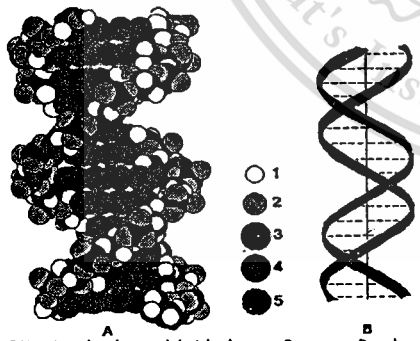
The concordance data show technical terms in their contexts of use. Figure 4.7 shows the concordance data for lexical collocations of the word *'dose'*. It consists of adjective (daily) and noun (dose).

4.6 Sample Entry for the Sample Dictionary

For the medical dictionary, medical terms are collected from the pharmaceutical corpus and prioritized on ground of overall frequency. In other words, the main factor in deciding whether or not to include a term as an entry in the dictionary is how often it occurs in the corpus because the frequency reflects the relative importance of words for the students. The selected entries are from technical vocabulary, and abbreviations.

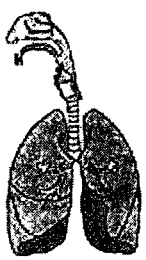
A	
acute /əˈkjuːt/ *adj.* **a :** characterized by sharpness or severity **b :** having a sudden onset, sharp rise, and short course. *A patient who survives the acute phase and is a symptomatic should be closely observed for at least 6 hrs.*	(โรค) ปัจจุบัน, เฉียบพลัน ***Origin:*** Acutus (L)
asthma /ˈæsmə/ *BrE.* /ˈæzmə/ ***n.*** **:** a condition often of allergic origin that is marked by continuous or paroxysmal labored breathing accompanied by wheezing, by a sense of constriction in the chest, and often by attacks of coughing or gasping. *The management of asthma should normally follow a stepwise programme, and response should be monitored clinically and by lung function tests.*	โรคหืด
C	
cardiac /ˈkɑːdi-æk/ *adj.* , ***n.*****1 a :** of, relating to, situated near, or acting on the heart **b :** of or relating to the cardiac of the stomach **2 :** of, relating to, or affected with heart disease. *In patients with cardiac failure, total plasma clearance will be reduced and lower dosages may be required.*	๑. เกี่ยวกับหัวใจ, เกี่ยวกับปากกระเพาะ ๒. ผู้ป่วยโรคหัวใจ ***Origin:*** Cardiacus [L] Kardiakos [Gr]

cancer /ˈkænsə/ ***n.*** : a malignant tumor of potentially unlimited growth that expands locally by invasion and systemically by metastasis. *Treatment of advanced breast **cancer** in women with natural or artificially induced postmenopausal status, who have previously been treated with antioestrogens.*	มะเร็ง
chronic /ˈkrɒnɪk/ ***adj.*** **1 a** : marked by long duration, by frequent recurrence over a long time, and often by slowly progressing seriousness : not acute **b** : suffering from a disease or ailment of long duration or frequent recurrence **2** : having a slow progressive course of indefinite duration -- used especially of degenerative invasive diseases, some infections, psychoses, and inflammations. *The adrenal reserve remains normal during **chronic** treatment as measured by a normal increment on a stimulation test.*	เรื้อรัง **Origin**: Chronicus [L] Chronos [Gr]
clinical /ˈklɪnɪkəl/ ***adj.*** : of, relating to, or conducted in or as if in a clinic: as **a** : involving or concerned with the direct observation and treatment of living **b** : of, relating to, based on, or characterized by observable and diagnosable symptoms of disease. *During **clinical** trials the following adverse events were reported.* **clinically** /ˈklɪnɪkəlɪ/ ***adv.***	เกี่ยวกับด้านการรักษา, เกี่ยวกับคลินิก
contraindication /ˌkɒntrəɪndɪˈkeɪʃən/ ***n.*** : something (as a symptom or condition) that makes a particular treatment or procedure inadvisable. *There are a number of **contraindications** in a person who has renal disease.*	ข้อห้ามใช้
D	
dilute /daɪˈluːt/ *v.* : to make thinner or more liquid by admixture *Taxol injection must be **diluted** prior to infusion.*	เจือ , ทำให้จืด, จาง ***Origin***: Dilutus [L]

disorder /dɪsˈɔːdə / ***n.*** **a:** a state of being badly prepared or organized for a particular task. **b:** an illness or disability which affects someone's mind or body. *There was evidence of a kidney **disorder**.*	๑. การเสียระเบียบ ๒. (จิตเวช) โรค, ความผิดปรกติ ***Origin***: Pref. Dis- + order: cf. Desordre[F]
diuretic /daɪˈretɪk/ ***adj., n.*** **:** tending to increase the excretion of urine **b:** a substance that increases the flow of urine. (adj.) *Consideration should be given to reducing the **diuretic** dose in patients with normal or low blood pressure who have been treated vigorously with diuretic or who are hyponatremic.* (n.) *Patients receiving **diuretics** should be adequately hydrated and be monitored for renal function prior to initiating treatment.*	๑. เกี่ยวกับขับปัสสาวะ ๒. ยาขับปัสสาวะ ***Origin***: Diouretikos [Gr]
DNA /diː en ˈeɪ/ ***n.*** **:** any of various nucleic acids that are usually the molecular basis of heredity, are localized especially in cell nuclei, and are constructed of a double helix held together by hydrogen bonds between purine and pyrimidine bases which project inward from two chains containing alternate links of deoxyribose and phosphate -- called also *deoxyribonucleic acid. The mechanism of action is related to its ability to bind to **DNA**.*  DNA: A molecular model: 1 hydrogen, 2 oxygen, 3 carbon in the helical phosphate ester chains, 4 carbon and nitrogen in the cross-linked purine and pyrimidine bases, 5 phosphorus; B double helix	ดีเอ็นเอ (กรดดีออกซีไรโบนิวคลิอิก)
dose /dəʊs / ***n.*** **a :** the measured quantity of a therapeutic agent to be taken at one time **b :** the quantity of radiation administered or absorbed.*The **dose** may be increased to 20 mg (20ml) after 1-2 weeks, depending on the clinical effect.*	ขนาดยา **Origin**: Dosis [Gr]

G	
gastrointestinal /ˌgæstrəʊɪnˈtestɪnəl/ ***adj.*** : of, relating to, or affecting both stomach and intestine. *The swallowed portion of an inhaled dose is absorbed from the **gastrointestinal** tract and undergoes considerable first-pass metabolism to the phenolic sulfate.*	๑.เกี่ยวกับกระเพาะและลำไส้ ๒. เกี่ยวกับทางเดินอาหาร
GI ***abbr.*** for gastrointestinal	
H	
hepatic /hɪˈpætɪk/ ***adj.*** : of, relating to, affecting, or associated with the liver. *No dose adjustments are required for patients with **hepatic** or renal insufficiency.*	เกี่ยวกับตับ **Origin**: Hepaticus [L] Hepatikos [Gr]
hypersensitivity /ˌhaɪpəsensɪˈtɪvɪty/ ***n.*** : excessively or abnormally sensitiv. *Patients who experience severe **hypersensitivity** reactions to Taxol should not be rechallenged with the drug.*	สภาพไวสูง
hypertension /ˌhaɪpəˈtenʃən/ ***n.*** : abnormally high arterial blood pressure : a systemic condition resulting from hypertension that is either symptomless or is accompanied by nervousness, dizziness, or headache. ***Hypertension** usually responds to a reduction in dose or discontinuation of therapy.*	ความดันเลือดสูง, ความดันสูง
hypotension /ˌhaɪpəʊˈtenʃən/ ***n.*** : abnormally low pressure of the blood -- called also *low blood pressure* : abnormally low pressure of the intraocular fluid. *The most likely event to occur is **hypotension** which shouid be treated by means of volume expansion.*	ความเลือดดันต่ำ ความดันต่ำ
I	
inhale /ɪnˈheɪl/ ***v.*** : to breathe or draw into the lungs; to inspire; as, to inhale air; opposed to exhale. *Doses to be **inhaled** up to 4 times in a 24-hr period.*	สูดหายใจ **Origin**: Inhalare [L]

impair /ɪmˈpeə/ *v.* : to damage or make worse by or as if by diminishing in some material respect. *Lowering the blood pressure may **impair** the patient's ability to concentrate and react.*	ทำให้เสียหาย
in vitro / ɪnˈvɪtrɔː/ *adv., adj.* : outside the living body and in an artificial environment. *The following **in vitro** data are available, but their clinical significance is unknown.*	นอกกาย
IV *abbr.* for **Intravenous**	
in vivo /ɪnˈvɪvəʊ/ *adv., adj* **1** : in the living body of a plant or animal **2** : in a real-life situation. *Enzyme inhibition by topotecan has not been evaluated **in vivo**.*	ในกาย
lactation /lækˈteɪʃən/ *n.* **1** : the secretion and yielding of milk by the mammary gland **2** : one complete period of lactation extending from about the time of parturition to weaning. *The preparation should not be used during **lactation**.*	๑. การหลั่งน้ำนม, การดูดนม ๒. ระยะหลั่งน้ำนม
M	
MAOI *abbr.* for **Monoamine Oxidase Inhibitor**	
metabolism /məˈtæbəlɪzm/ *n.* : the sum of the processes in the buildup and destruction of protoplasm ; *specifically* : the chemical changes in living cells by which energy is provided for vital processes and activities and new material is assimilated. *The disturbed cellular **metabolism** is normalized.*	กระบวนการสร้างและสลาย, เมแทบอลิซึม **Origin**: Metaballein [Gr]
N	
NSAID *abbr.* for : a nonsteroidal anti-inflammatory drug	
O	
oral /ˈɔːrəl/ *adj.* **a** : of, relating to, or involving the mouth **b** : given or taken through or by way of the mouth. *The infusion is continued for at least 5 days, to allow impregnation by the **oral** administration.*	เกี่ยวกับปาก, ทางปาก

overdosage /ˌəʊvəˈdəʊsɪdʒ/ ***n.*** **1 :** the administration or taking of an excessive dose **2 :** the condition of being overdosed. *The dose can be increased until the following signs of **overdosage** are observal.*	๑. การให้ยาเกิน ๒. ขนาดยาเกิน
P	
pharmacokinetics /ˌfɑːməkəʊkɪˈnɛtɪks/ ***n.*** **1 :** the study of the bodily absorption, distribution, metabolism, and excretion of drugs **2 :** the characteristic interactions of a drug and the body in terms of its absorption, distribution, metabolism, and excretion. *The **pharmacokinetics** of capecitabine have been evaluated over a dose range of 502-3514 mg/m2/day.*	เภสัชจลนศาสตร์
plasma /ˈplæzə/ ***n.*** **:** the fluid part especially of blood, lymph, or milk that is distinguished from suspended material. *The major metabolite identified in human **plasma** is O-desmethyl gefitinib.*	พลาสมา
pregnancy /ˈpregnənsi/ ***n.*** **1 :** the condition of being pregnant **2 :** an instance of being pregnant. *Administration of Voltaren ampoules during **pregnancy** and lactation is not recommended.*	การตั้งครรภ์
R	
respiratory /ˈrespɪrɪətərɪ/ ***adj.*** **:** of or relating to respiration, serving for or functioning in respiration nerves. *Special care is needed in patients with active or quiescent pulmonary tuberculosis or fungal, bacterial or viral infections of the **respiratory** system.*  Anatomy of the respiratory system.	เกี่ยวกับหายใจ

V	
vial /'vaɪəl/ ***n.*** : a small closed or closable vessel especially for liquids -- called also *phial. After adding the sterile solvent, shake the vial until the drug has completely dissolved.* 	ขวดเล็ก

BIBLIOGRAPHY

Anderson, K. (ed). 2002. **Mosby's Medical, Nursing, & Allied Health Dictionary, 6th Ed.** Mosby Inc.

Andrews, S. 1981. Specificity in Language Course Design. In **English Language Research Journal**, ed. Chris Kennedy, 12-13. Bangkok : Library Faculty of Science Mahidol University.

Barnhart, R.K. (Ed.) 1995. **Barnhart Abbreviations Dictionary.** United States of America: John Wiley & Sons, Inc.

Baumand, J. and Culigan, B. 1995. **The General Service List.** [Online]. Available: http://Jbauman.com/gsl.html

Benson, M., Benson, E. & Ilson, R. 1986. **The BBI Combinatory Dictionary of English.** John Benjamins Publishing Company.

Bergenjoltz, H. and S. Trap. 1995. **Manual of Specialised of Lexicography: The Preparation of Specialised Dictionaries.** Amsterdam/Philadelphia. John Benjamin's Publishing Company.

Biber, D. et al. 1998. **Corpus Linguistics: Investigation Language Structure and Use.** Cambridge : Cambridge University Press.

Biber, D. et al. 1999. **Longman Grammar of Spoken and Written English.** Harlow: Pearson Education Limited.

Boonya-udomsart, D. 1976. **An Analysis of General English Vocabulary and Frequency in an Economics Textbook.** M.Ed. Thesis, Department of Education , Graduate School, Srinakarinwirot University Prasanmitr.

Boswood, T.(Ed.) 1997. **New Way of Using Computers in Language Teaching.** Illinois: TESOL Inc., pp. 253-275.

Brasnett, C. 1976. **English for Medical Students.** London : Methuen Education.

Coxhead, A. 2000. A new academic word list. **TESOL Quarterly, 34,** 213-238.

Dudley, E. & John, M. 1998. **Developments in English for Specific Purposes: a Multi-disciplinary Approach.** Cambridge : Cambridge University Press.

Flowerdew, J. 1993. Variation across Speech and Writing in Biology. Hong Kong: Public Lecture Given at City Polytechnic of Hong Kong.

Glendinning, Eric, and Beverly Holmstrom. 1995. **English in Medicine : a Course in Communication Skills.** 3d ed. Cambridge : Cambridge University Press.

Godman, A. and E M F. Payne. 1979. **Longman Dictionary of Scientific Usage.** Harlow: Longman.

Hoffmann, L. 1995. **Kommunikationsmittel faxhsprache.** Tubingen: Gunter Narr Verleg.

Hunston, S. 2002. **Corpora in Applied Linguistics.** Cambridge: Cambridge University Press.

Hutchison, T. and Waters, A. 1986. **English for Special Purposes: A Learning Centred Approach.** Cambridge: Cambridge University Press.

Jabbour, G. N. 1998. **Corpus Linguistics, Contextual Collocation and ESP Syllabus Creation: a Text Analysis Approach to the Study of Medical Research Articles.** Ph.D. Thesis. University of Birmingham,
United Kingdom. . Available: http://thailis.uni.net.th/dao/detail.nsp

Jeffery, G. 1953. Foreword in **General Service List of English Words.** London: Longman, Gree & Co.

Jutharat Prawanpath. 1984. **A study of general vocabulary knowledge of the pre-medical science students at Mahidol University.** Thesis. Graduate School, Mahidol University.

Keawphanngam, C. 2002. **A Corpus Analysis of Psychology Texts as a Basis for the Development of Teaching Materials in English for Academic Purposes.** M.A. Thesis. Graduate School, Mahidol University.

Kerr, L. 1979. English for Special Purposes. In **English for Specific Purposes,** ed. Susan Holden,11-12. Oxford : Modern English Publication.

Krishnamurthy, R. 2000. Classroom Cornucopia: the COBUILD Dictionary and the Bank of English Corpus. Available: http://www.cobuild.collins.co.uk

Kucera, H. and Francis, W.N. 1967. Computational Analysis of present-Day American English. Providence, RI: Brown University Press

Landau, S. 2001. **Dictionaries: the Art and Craft of Lexicography.** Cambridge: Cambridge University Press.

Mackay, Ronal, and A. Mountford. 1978. **English for Specific Purposes : Applied Linguistics and Language study**. 2d ed. London : Longman

Maclean, J. 1975. **English in Basic Medical Science.** London: Oxford University Press.

Maher, J. 1986. *'English for Medical Purposes : State of the Art Article'*, **Language Teaching,** 19 : 112-145.

Maher, J. and Ann Arbor. 1992. *'Review of International Medical Communication in English by Francoise Salager-Meyer'*, **English for Specific Purposes,** 11 : 269-271.

Mallikamas, P. 1999. Applications of Corpora in Language Teaching. **ThaiTESOL Bulletin,**12(1), pp.1-17.

Meijs, W. 1996. Linguistic Corpora and Lexicography. **Annual Review of Applied Linguistics,**16, pp.99-114.

Ministry of Public Health. 2000. Public Health Statistics 2000. Bangkok: Ministry of Public Health.

Nation, I.S.P. 2001. **Learning vocabulary in another language**. Cambridge: Cambridge University Press.

Opitz, K. 1983. Special-purpose Lexicography: Dictionaries for Technical Use. In R.R.K. Hartmant (Ed) **Dictionaries and their Users, Exeter Linguistic Studies 4.** Exeter: University of Exeter: 89-95.

Paiva, D. 1999. **Investigating style in a corpus of pharmaceutical leaflets: results of a factor analysis.** [Online]. Avialable:www.itri.brighton.ac.uk/~Daniel.Paiva/acl2000student.finalversion.pdf

Pratt, W. et al. 2002. **Changing the paradigm for medical English language Teaching.** [Online]. Available: http://www.usingenglish.com/articles/medical-english-language-teaching.html

Pearson, J. 1998. **Terms in Context.** Amsterdam: John Benjamin's Publishing company.

Renouf, A.J. 1978. 'Corpus Development'. In J.M. Sinclair (ed.) **Looking Up: An Account of the COBUILD Project in Lexical Computing.** London: Collins.

Richard, J. 1999. **Dictionary of Language Teaching & Applied Linguistics.** London. Longman.

Robinson, P. 1991. **ESP Today: A Practitioner's Guide.** Prentice Hall International.

Scott, M. 1996. **WordSmith Tools (Version 3.0).** [Computer software]. Oxford: Oxford University Press.

Simon, M. 1996. Linguistic corpora and language teaching. **Annual Review of Applied Linguistics, 16,** pp. 82-199.

Sinclair, J. (Ed). 1987. **Looking Up: Account of the COBUILD Project**. Scotprint Ltd., Musselburgh.

Swales, J. M. 2000. Language for Specific Purposes. **Annual Review of Applied Linguistics(20),** pp.59-76.

The Thai Royal Institute. 1996. **Medical Terminology**.[DC-ROM]

Thorndike, E. and Lorge, I. 1944. **the Teacher's Wordbook of 30,000 Words**. New York: Teacher's College.

Thurstun, J. and Candlin, C. 1998. Concordancing and Teaching of the Vocabulary of Academic English. **English for Specific Purposes,17**(3), pp.267-280.

Tribble, C. and Jones, G. 1990. **Concordance in the classroom**. London. Longman.

Trimble, L. 1985. **English for Science and Technology**. Cambridge: Cambridge University Press.

Trimble, R.M.T. and Trimble, L. 1978. 'The Development of EFL Materials for Occupational English': In R.M.T. Trimble, L. Trimble and K. Drobonic (eds.). **English for Specific Purposes, Science and Technology.** English Language Institute. Oregon State University: 74-132.

West, M. 1953. **General Service List of English Words**. London: Longman, Green & Co.

World Health Organization (WHO). **WHO Ethical Criteria for Medical Drug Promotion**. Report by Director-General; Nov 15, 1993.

Widdowson, H.G. 1998. *Communication and Community : the Pragmatics of ESP*, **ESP(s) An International Journal** 17,1 : 3-13.

APPENDICES

APPENDIX A

The General Service List

(Bauman amd Culligan, 1995)

This version of the General Service List is the list of 2,248 words created by John Bauman and Brent Culligan in 1995. This list includes all the 2,000 word from original General Service List of West (1953), plus 248 more, ranked and presented in frequency order based on the Brown Corpus, 1982.

1.	the	39.	can	77.	think	115.	nation
2.	be	40.	more	78.	most	116.	hand
3.	of	41.	if	79.	even	117.	old
4.	and	42.	no	80.	find	118.	life
5.	a	43.	man	81.	day	119.	tell
6.	to	44.	out	82.	also	120.	write
7.	in	45.	other	83.	after	121.	become
8.	he	46.	so	84.	way	122.	here
9.	have	47.	what	85.	many	123.	show
10.	it	48.	time	86.	must	124.	house
11.	that	49.	up	87.	look	125.	both
12.	for	50.	go	88.	before	126.	between
13.	they	51.	about	89.	great	127.	need
14.	I	52.	than	90.	back	128.	mean
15.	with	53.	into	91.	through	129.	call
16.	as	54.	could	92.	long	130.	develop
17.	not	55.	state	93.	where	131.	under
18.	on	56.	only	94.	much	132.	last
19.	she	57.	new	95.	should	133.	right
20.	at	58.	year	96.	well	134.	move
21.	by	59.	some	97.	people	135.	thing
22.	this	60.	take	98.	down	136.	general
23.	we	61.	come	99.	own	137.	school
24.	you	62.	these	100.	just	138.	never
25.	do	63.	know	101.	because	139.	same
26.	but	64.	see	102.	good	140.	another
27.	from	65.	use	103.	each	141.	begin
28.	or	66.	get	104.	those	142.	while
29.	which	67.	like	105.	feel	143.	number
30.	one	68.	then	106.	seem	144.	part
31.	would	69.	first	107.	how	145.	turn
32.	all	70.	any	108.	high	146.	real
33.	will	71.	work	109.	too	147.	leave
34.	there	72.	now	110.	place	148.	might
35.	say	73.	may	111.	little	149.	want
36.	who	74.	such	112.	world	150.	point
37.	make	75.	give	113.	very	151.	form
38.	when	76.	over	114.	still	152.	off

153. child
154. few
155. small
156. since
157. against
158. ask
159. late
160. home
161. interest
162. large
163. person
164. end
165. open
166. public
167. follow
168. during
169. present
170. without
171. again
172. hold
173. govern
174. around
175. possible
176. head
177. consider
178. word
179. program
180. problem
181. however
182. lead
183. system
184. set
185. order
186. eye
187. plan
188. run
189. keep
190. face
191. fact
192. group
193. play
194. stand
195. increase
196. early
197. course
198. change
199. help
200. line
201. city
202. put
203. close
204. case
205. force
206. meet
207. once
208. water
209. upon
210. war
211. build
212. hear
213. light
214. unite
215. live
216. every
217. country
218. bring
219. center
220. let
221. side
222. try
223. provide
224. continue
225. name
226. certain
227. power
228. pay
229. result
230. question
231. study
232. woman
233. member
234. until
235. far
236. night
237. always
238. service
239. away
240. report
241. something
242. company
243. week
244. church
245. toward
246. start
247. social
248. room
249. figure
250. nature
251. though
252. young
253. less
254. enough
255. almost
256. read
257. include
258. president
259. nothing
260. yet
261. better
262. big
263. boy
264. cost
265. business
266. value
267. second
268. why
269. clear
270. expect
271. family
272. complete
273. act
274. sense
275. mind
276. experience
277. art
278. next
279. near
280. direct
281. car
282. law
283. industry
284. important
285. girl
286. god
287. several
288. matter
289. usual
290. rather
291. per
292. often
293. kind
294. among
295. white
296. reason
297. action
298. return
299. foot
300. care
301. simple
302. within
303. love
304. human
305. along
306. appear
307. doctor
308. believe
309. speak
310. active
311. student
312. month
313. drive
314. concern
315. best
316. door
317. hope
318. example
319. inform
320. body
321. ever
322. least
323. probable
324. understand
325. reach
326. effect
327. different
328. idea
329. whole
330. control
331. condition
332. field
333. pass
334. fall
335. note
336. special
337. talk
338. particular
339. today
340. measure
341. walk
342. teach
343. low
344. hour
345. type
346. carry
347. rate
348. remain
349. full
350. street
351. easy
352. although
353. record
354. sit
355. determine
356. level
357. local
358. sure
359. receive
360. thus
361. moment
362. spirit
363. train
364. college
365. religion
366. perhaps
367. music
368. grow

369. free
370. cause
371. serve
372. age
373. book
374. board
375. recent
376. sound
377. office
378. cut
379. step
380. class
381. true
382. history
383. position
384. above
385. strong
386. friend
387. necessary
388. add
389. court
390. deal
391. tax
392. support
393. party
394. whether
395. either
396. land
397. material
398. happen
399. education
400. death
401. agree
402. arm
403. mother
404. across
405. quite
406. anything
407. town
408. past
409. view
410. society
411. manage
412. answer
413. break
414. organize
415. half
416. fire
417. lose
418. money
419. stop
420. actual
421. already
422. effort
423. wait
424. department
425. able
426. political
427. learn
428. voice
429. air
430. together
431. shall
432. cover
433. common
434. subject
435. draw
436. short
437. wife
438. treat
439. limit
440. road
441. letter
442. color
443. behind
444. produce
445. send
446. term
447. total
448. university
449. rise
450. century
451. success
452. minute
453. remember
454. purpose
455. test
456. fight
457. watch
458. situation
459. south
460. ago
461. difference
462. stage
463. father
464. table
465. rest
466. bear
467. entire
468. market
469. prepare
470. explain
471. offer
472. plant
473. charge
474. ground
475. west
476. picture
477. hard
478. front
479. lie
480. modern
481. dark
482. surface
483. rule
484. regard
485. dance
486. peace
487. observe
488. future
489. wall
490. farm
491. claim
492. firm
493. operation
494. further
495. pressure
496. property
497. morning
498. amount
499. top
500. outside
501. piece
502. sometimes
503. beauty
504. trade
505. fear
506. demand
507. wonder
508. list
509. accept
510. judge
511. paint
512. mile
513. soon
514. responsible
515. allow
516. secretary
517. heart
518. union
519. slow
520. island
521. enter
522. drink
523. story
524. experiment
525. stay
526. paper
527. space
528. apply
529. decide
530. share
531. desire
532. spend
533. sign
534. therefore
535. various
536. visit
537. supply
538. officer
539. doubt
540. private
541. immediate
542. wish
543. contain
544. feed
545. raise
546. describe
547. ready
548. horse
549. son
550. exist
551. north
552. suggest
553. station
554. effective
555. food
556. deep
557. wide
558. alone
559. character
560. English
561. happy
562. critic
563. unit
564. product
565. respect
566. drop
567. nor
568. fill
569. cold
570. represent
571. sudden
572. basic
573. kill
574. fine
575. trouble
576. mark
577. single
578. press
579. heavy
580. attempt
581. origin
582. standard
583. everything
584. committee

585. moral
586. black
587. red
588. bad
589. earth
590. accord
591. else
592. mere
593. die
594. remark
595. basis
596. except
597. equal
598. east
599. event
600. employ
601. defense
602. smile
603. river
604. improve
605. game
606. detail
607. account
608. cent
609. sort
610. reduce
611. club
612. buy
613. attention
614. ship
615. decision
616. wear
617. inside
618. win
619. suppose
620. ride
621. operate
622. realize
623. sale
624. choose
625. park
626. square
627. vote
628. price
629. district
630. dead
631. foreign
632. window
633. beyond
634. direction
635. strike
636. instead
637. trial
638. practice
639. catch
640. opportunity
641. likely
642. recognize
643. permit
644. serious
645. attack
646. floor
647. association
648. spring
649. lot
650. stock
651. lack
652. hair
653. science
654. relation
655. profession
656. pattern
657. quick
658. medical
659. influence
660. occasion
661. machine
662. compare
663. husband
664. blue
665. international
666. fair
667. especially
668. indeed
669. imagine
670. surprise
671. average
672. official
673. temperature
674. difficult
675. sing
676. hit
677. tree
678. race
679. police
680. touch
681. relative
682. throw
683. quality
684. former
685. pull
686. chance
687. prove
688. argue
689. settle
690. growth
691. date
692. heat
693. save
694. performance
695. count
696. production
697. listen
698. main
699. pick
700. size
701. cool
702. army
703. patient
704. combine
705. summer
706. hall
707. slight
708. command
709. enjoy
710. length
711. proper
712. express
713. health
714. chief
715. evening
716. store
717. language
718. degree
719. lay
720. current
721. gun
722. dog
723. hotel
724. strange
725. separate
726. boat
727. fail
728. clean
729. dress
730. anyone
731. gain
732. pain
733. object
734. knowledge
735. depend
736. relate
737. below
738. dollar
739. advance
740. shape
741. arrange
742. population
743. yes
744. sell
745. mention
746. dry
747. check
748. poet
749. sleep
750. join
751. hot
752. bed
753. electric
754. dream
755. due
756. season
757. manner
758. fit
759. left
760. progress
761. neither
762. strength
763. notice
764. finish
765. opinion
766. bill
767. western
768. truth
769. wrong
770. travel
771. suit
772. bank
773. exact
774. honor
775. brother
776. quiet
777. marry
778. corner
779. handle
780. danger
781. hospital
782. pool
783. promise
784. blood
785. shoot
786. scene
787. literature
788. arrive
789. film
790. base
791. freedom
792. bar
793. maybe
794. hang
795. suffer
796. manufacture
797. frequent
798. rock
799. loss
800. burn

801. sun
802. audience
803. essential
804. glass
805. prevent
806. poem
807. poor
808. inch
809. song
810. skill
811. post
812. popular
813. radio
814. animal
815. conscious
816. worth
817. eat
818. election
819. faith
820. wave
821. murder
822. model
823. forget
824. extend
825. edge
826. distance
827. memory
828. recommend
829. division
830. staff
831. leg
832. discussion
833. address
834. fly
835. dependent
836. ball
837. shake
838. frame
839. extreme
840. engineer
841. thick
842. comfort
843. latter
844. camp
845. oil
846. discover
847. examine
848. difficulty
849. tooth
850. middle
851. choice
852. refer
853. enemy
854. practical
855. marriage
856. bridge
857. declare
858. lady
859. cross
860. daily
861. afternoon
862. attend
863. director
864. balance
865. wash
866. capital
867. speed
868. block
869. citizen
870. mouth
871. hill
872. green
873. please
874. motor
875. agency
876. encourage
877. governor
878. worry
879. affair
880. shoulder
881. bright
882. mass
883. sample
884. pretty
885. repeat
886. roll
887. push
888. trip
889. council
890. clothe
891. parent
892. forward
893. sharp
894. straight
895. gas
896. weight
897. discuss
898. fix
899. load
900. master
901. whatever
902. round
903. rapid
904. laugh
905. finger
906. spot
907. propose
908. shop
909. broad
910. replace
911. reply
912. extent
913. lock
914. employee
915. ahead
916. sight
917. spread
918. wind
919. approve
920. destroy
921. none
922. pound
923. fame
924. importance
925. reflect
926. advantage
927. match
928. regular
929. wage
930. refuse
931. existence
932. hardly
933. perform
934. title
935. tend
936. exercise
937. thin
938. coat
939. bit
940. mountain
941. youth
942. behavior
943. newspaper
944. secret
945. ability
946. sea
947. soft
948. justice
949. reasonable
950. circle
951. solid
952. page
953. weapon
954. fast
955. representative
956. search
957. pure
958. escape
959. crowd
960. stick
961. telephone
962. avoid
963. garden
964. favor
965. news
966. unless
967. dinner
968. someone
969. signal
970. yard
971. ideal
972. warm
973. miss
974. shelter
975. soldier
976. article
977. cry
978. captain
979. familiar
980. seat
981. guest
982. weak
983. excite
984. king
985. everyone
986. wine
987. hole
988. duty
989. beat
990. perfect
991. bottom
992. compose
993. battle
994. expense
995. cattle
996. flow
997. kitchen
998. dust
999. bottle
1000. admit
1001. tear
1002. tire
1003. expression
1004. exception
1005. application
1006. belong
1007. rich
1008. failure
1009. struggle
1010. instrument
1011. variety
1012. narrow
1013. theater
1014. collection
1015. rain
1016. review

1017. preserve
1018. leadership
1019. clay
1020. daughter
1021. fellow
1022. swing
1023. thank
1024. library
1025. fat
1026. reserve
1027. tour
1028. nice
1029. warn
1030. ring
1031. bitter
1032. chair
1033. yesterday
1034. scientific
1035. flower
1036. wheel
1037. solution
1038. aim
1039. gather
1040. invite
1041. moreover
1042. fresh
1043. forest
1044. winter
1045. box
1046. belief
1047. ordinary
1048. impossible
1049. print
1050. gray
1051. taste
1052. lip
1053. speech
1054. reference
1055. stain
1056. connection
1057. otherwise
1058. stretch
1059. knife
1060. village
1061. blow
1062. mistake
1063. sweet
1064. shout
1065. divide
1066. guard
1067. worse
1068. exchange
1069. rare
1070. commercial
1071. request
1072. appoint
1073. agent
1074. dependence
1075. bird
1076. wild
1077. motion
1078. guess
1079. neighbor
1080. seed
1081. fashion
1082. loan
1083. correct
1084. plain
1085. mail
1086. retire
1087. opposite
1088. prefer
1089. safe
1090. evil
1091. double
1092. wood
1093. empty
1094. baby
1095. advise
1096. content
1097. sport
1098. lift
1099. literary
1100. curious
1101. tie
1102. flat
1103. message
1104. neck
1105. hate
1106. dirt
1107. delight
1108. trust
1109. nobody
1110. valley
1111. tool
1112. presence
1113. cook
1114. railroad
1115. minister
1116. coffee
1117. brush
1118. beside
1119. collect
1120. guide
1121. luck
1122. profit
1123. lord
1124. everybody
1125. prison
1126. cloud
1127. slave
1128. chairman
1129. soil
1130. distinguish
1131. introduce
1132. urge
1133. blind
1134. arise
1135. upper
1136. curve
1137. membership
1138. key
1139. entertain
1140. soul
1141. neighborhood
1142. friendly
1143. pair
1144. stone
1145. lean
1146. protect
1147. advertise
1148. mystery
1149. welcome
1150. knee
1151. jump
1152. snake
1153. stream
1154. avenue
1155. brown
1156. disease
1157. hat
1158. excellent
1159. formal
1160. snow
1161. sheet
1162. somehow
1163. unity
1164. sky
1165. rough
1166. smooth
1167. weather
1168. steady
1169. threaten
1170. depth
1171. oppose
1172. deliver
1173. ancient
1174. pray
1175. adopt
1176. birth
1177. appearance
1178. universe
1179. busy
1180. hurry
1181. coast
1182. forth
1183. smell
1184. furnish
1185. female
1186. hide
1187. wire
1188. proposal
1189. ought
1190. victory
1191. quarter
1192. engine
1193. customer
1194. waste
1195. fool
1196. intend
1197. intention
1198. desk
1199. politics
1200. passage
1201. lawyer
1202. root
1203. climb
1204. metal
1205. gradual
1206. hunt
1207. protection
1208. satisfy
1209. roof
1210. branch
1211. pleasure
1212. witness
1213. loose
1214. nose
1215. mine
1216. band
1217. aside
1218. risk
1219. tomorrow
1220. remind
1221. ear
1222. fish
1223. shore
1224. operator
1225. civilize
1226. being
1227. silent
1228. screen
1229. bind
1230. earn
1231. pack
1232. colony

1233. besides
1234. slip
1235. cousin
1236. scale
1237. relief
1238. explore
1239. stem
1240. brain
1241. musician
1242. defend
1243. bend
1244. somebody
1245. shadow
1246. mix
1247. smoke
1248. description
1249. fruit
1250. guilt
1251. yield
1252. sensitive
1253. salt
1254. pale
1255. sweep
1256. completion
1257. throat
1258. agriculture
1259. admire
1260. gentle
1261. dozen
1262. particle
1263. pleasant
1264. bay
1265. cup
1266. competition
1267. moon
1268. terrible
1269. strip
1270. mechanic
1271. shock
1272. conversation
1273. angle
1274. tall
1275. plenty
1276. star
1277. yellow
1278. sick
1279. thorough
1280. absolute
1281. succeed
1282. surround
1283. proud
1284. dear
1285. card
1286. lake
1287. breath
1288. afraid
1289. silence
1290. onto
1291. shoe
1292. somewhere
1293. chain
1294. slide
1295. copy
1296. machinery
1297. wake
1298. severe
1299. pocket
1300. bone
1301. honest
1302. freeze
1303. dictionary
1304. calm
1305. swim
1306. ice
1307. male
1308. skin
1309. crack
1310. rush
1311. wet
1312. meat
1313. commerce
1314. joint
1315. gift
1316. host
1317. suspect
1318. path
1319. uncle
1320. afford
1321. instant
1322. satisfactory
1323. height
1324. track
1325. confidence
1326. grass
1327. suggestion
1328. favorite
1329. breakfast
1330. apart
1331. chest
1332. entrance
1333. march
1334. sink
1335. northern
1336. iron
1337. alive
1338. ill
1339. bag
1340. disturb
1341. native
1342. bedroom
1343. violent
1344. beneath
1345. pause
1346. tough
1347. substance
1348. threat
1349. charm
1350. absence
1351. factory
1352. spite
1353. meal
1354. universal
1355. accident
1356. highway
1357. sentence
1358. liberty
1359. wise
1360. noise
1361. discovery
1362. tube
1363. flash
1364. twist
1365. fence
1366. childhood
1367. joy
1368. sister
1369. sad
1370. efficiency
1371. disappear
1372. defeat
1373. extensive
1374. rent
1375. comparison
1376. possess
1377. grace
1378. flesh
1379. liquid
1380. scientist
1381. ease
1382. heaven
1383. milk
1384. sympathy
1385. rank
1386. restaurant
1387. frequency
1388. angry
1389. shade
1390. accuse
1391. necessity
1392. knock
1393. loud
1394. permanent
1395. row
1396. lovely
1397. confuse
1398. gold
1399. frighten
1400. solve
1401. grave
1402. salary
1403. photograph
1404. advice
1405. abroad
1406. wound
1407. virtue
1408. dare
1409. queen
1410. extra
1411. attract
1412. numerous
1413. pink
1414. gate
1415. expensive
1416. shut
1417. chicken
1418. forgive
1419. holy
1420. wooden
1421. prompt
1422. crime
1423. sorry
1424. republic
1425. anger
1426. visitor
1427. pile
1428. violence
1429. steel
1430. wing
1431. stair
1432. partner
1433. delay
1434. gentleman
1435. pour
1436. confusion
1437. damage
1438. kick
1439. safety
1440. burst
1441. network
1442. resistance
1443. screw
1444. pride
1445. till
1446. hire
1447. verb
1448. preach

1449. clerk
1450. everywhere
1451. anyway
1452. fan
1453. connect
1454. egg
1455. efficient
1456. grain
1457. calculate
1458. drag
1459. opposition
1460. worship
1461. arrest
1462. discipline
1463. string
1464. harbor
1465. camera
1466. mechanism
1467. cow
1468. grand
1469. funny
1470. insurance
1471. reduction
1472. strict
1473. lesson
1474. tight
1475. sand
1476. plate
1477. qualify
1478. elsewhere
1479. mad
1480. interference
1481. pupil
1482. fold
1483. royal
1484. valuable
1485. whisper
1486. anybody
1487. hurt
1488. excess
1489. quantity
1490. fun
1491. mud
1492. extension
1493. recognition
1494. kiss
1495. crop
1496. sail
1497. attractive
1498. habit
1499. relieve
1500. wisdom
1501. persuade
1502. certainty
1503. cloth
1504. eager
1505. deserve
1506. sympathetic
1507. cure
1508. trap
1509. puzzle
1510. powder
1511. raw
1512. mankind
1513. glad
1514. blame
1515. whenever
1516. anxiety
1517. bus
1518. tremble
1519. sacred
1520. fortunate
1521. glory
1522. golden
1523. neat
1524. weekend
1525. treasury
1526. overcome
1527. cat
1528. sacrifice
1529. complain
1530. elect
1531. roar
1532. sake
1533. temple
1534. self
1535. compete
1536. nurse
1537. stuff
1538. stomach
1539. peculiar
1540. repair
1541. storm
1542. ton
1543. desert
1544. allowance
1545. servant
1546. hunger
1547. conscience
1548. bread
1549. crash
1550. tip
1551. strengthen
1552. proof
1553. generous
1554. sir
1555. tonight
1556. whip
1557. tongue
1558. mill
1559. merchant
1560. coal
1561. ruin
1562. introduction
1563. courage
1564. actor
1565. belt
1566. stir
1567. package
1568. punish
1569. reflection
1570. breathe
1571. anywhere
1572. amuse
1573. dull
1574. fate
1575. net
1576. fellowship
1577. fault
1578. furniture
1579. beam
1580. pencil
1581. border
1582. disappoint
1583. flame
1584. joke
1585. bless
1586. corn
1587. shell
1588. tempt
1589. supper
1590. destruction
1591. dive
1592. anxious
1593. shine
1594. cheap
1595. dish
1596. distant
1597. greet
1598. flood
1599. excuse
1600. insect
1601. ocean
1602. ceremony
1603. decrease
1604. prize
1605. harm
1606. insure
1607. verse
1608. pot
1609. sincere
1610. cotton
1611. leaf
1612. rub
1613. medicine
1614. stroke
1615. bite
1616. lung
1617. lonely
1618. admission
1619. stupid
1620. scratch
1621. composition
1622. broadcast
1623. drum
1624. resist
1625. neglect
1626. absent
1627. passenger
1628. adventure
1629. beg
1630. pipe
1631. beard
1632. bold
1633. meanwhile
1634. devil
1635. cheer
1636. nut
1637. split
1638. melt
1639. swear
1640. sugar
1641. bury
1642. wipe
1643. faint
1644. creature
1645. tail
1646. wealth
1647. earnest
1648. translate
1649. suspicion
1650. noble
1651. inquiry
1652. journey
1653. hesitate
1654. extraordinary
1655. borrow
1656. owe
1657. funeral
1658. ambition
1659. mixture
1660. slope
1661. criminal
1662. seldom
1663. map
1664. spin

1665. praise
1666. spare
1667. plow
1668. telegraph
1669. barrel
1670. straighten
1671. scarce
1672. lunch
1673. slavery
1674. creep
1675. sweat
1676. gay
1677. stiff
1678. brave
1679. seize
1680. convenient
1681. horizon
1682. moderate
1683. complicate
1684. dig
1685. curse
1686. weigh
1687. priest
1688. excessive
1689. quarrel
1690. widow
1691. modest
1692. dine
1693. politician
1694. custom
1695. educate
1696. salesman
1697. nail
1698. tap
1699. eastern
1700. possession
1701. satisfaction
1702. behave
1703. mercy
1704. scatter
1705. objection
1706. silver
1707. tent
1708. saddle
1709. wrap
1710. nest
1711. grind
1712. spell
1713. plaster
1714. arch
1715. swell
1716. friendship
1717. bath
1718. bundle
1719. grateful
1720. crown
1721. boundary
1722. nowhere
1723. asleep
1724. clock
1725. boil
1726. altogether
1727. lend
1728. holiday
1729. precious
1730. wander
1731. ugly
1732. reputation
1733. ticket
1734. pretend
1735. dismiss
1736. delicate
1737. despair
1738. awake
1739. tea
1740. FALSE
1741. fortune
1742. cap
1743. thread
1744. haste
1745. bare
1746. shirt
1747. bargain
1748. leather
1749. rail
1750. butter
1751. dot
1752. inquire
1753. warmth
1754. decisive
1755. vessel
1756. pity
1757. steam
1758. pin
1759. bound
1760. companion
1761. toe
1762. reward
1763. forbid
1764. wherever
1765. tower
1766. bathe
1767. lodge
1768. swallow
1769. multiply
1770. bow
1771. kingdom
1772. garage
1773. permission
1774. pump
1775. prevention
1776. urgent
1777. aunt
1778. zero
1779. idle
1780. fever
1781. Christmas
1782. regret
1783. jaw
1784. soap
1785. pronounce
1786. empire
1787. bowl
1788. outline
1789. organ
1790. imitation
1791. caution
1792. mineral
1793. disagree
1794. blade
1795. trick
1796. treasure
1797. immense
1798. convenience
1799. disapprove
1800. destructive
1801. fork
1802. noon
1803. ownership
1804. tune
1805. polish
1806. poison
1807. shame
1808. loyalty
1809. cottage
1810. astonish
1811. shave
1812. feather
1813. sauce
1814. lid
1815. debt
1816. fade
1817. confess
1818. classification
1819. descend
1820. cape
1821. mild
1822. clever
1823. envelope
1824. invention
1825. sheep
1826. splendid
1827. stamp
1828. float
1829. brick
1830. rice
1831. businessman
1832. backward
1833. qualification
1834. artificial
1835. attraction
1836. lamp
1837. curl
1838. shower
1839. elder
1840. bunch
1841. bell
1842. steer
1843. flavor
1844. spit
1845. rob
1846. cream
1847. interrupt
1848. pen
1849. weave
1850. orange
1851. rescue
1852. crush
1853. humble
1854. fancy
1855. decay
1856. polite
1857. tribe
1858. bleed
1859. coin
1860. fond
1861. autumn
1862. classify
1863. omit
1864. loyal
1865. needle
1866. lessen
1867. complaint
1868. pad
1869. steep
1870. skirt
1871. curtain
1872. calculation
1873. laughter
1874. solemn
1875. grease
1876. interfere
1877. explode
1878. fasten
1879. flag
1880. resign

1881. postpone
1882. patience
1883. boast
1884. rope
1885. envy
1886. airplane
1887. rid
1888. shield
1889. veil
1890. kneel
1891. tray
1892. explosive
1893. brass
1894. taxi
1895. wax
1896. duck
1897. button
1898. invent
1899. remedy
1900. bush
1901. thunder
1902. weaken
1903. poverty
1904. scrape
1905. arrow
1906. tender
1907. cruel
1908. soften
1909. mouse
1910. hay
1911. anyhow
1912. alike
1913. circular
1914. juice
1915. shelf
1916. bake
1917. hatred
1918. cautious
1919. basket
1920. wreck
1921. width
1922. confident
1923. log
1924. heap
1925. suck
1926. ladder
1927. gap
1928. obey
1929. hut
1930. axe
1931. translation
1932. collar
1933. delivery
1934. reproduce
1935. confession
1936. pan
1937. prejudice
1938. voyage
1939. tobacco
1940. simplicity
1941. paste
1942. cake
1943. elephant
1944. ribbon
1945. harvest
1946. ashamed
1947. cave
1948. customary
1949. thief
1950. damp
1951. sew
1952. rust
1953. separation
1954. waiter
1955. pet
1956. straw
1957. upset
1958. towel
1959. refresh
1960. essence
1961. fur
1962. ambitious
1963. defendant
1964. daylight
1965. dip
1966. suspicious
1967. imaginary
1968. ash
1969. carriage
1970. educator
1971. saw
1972. stove
1973. rubber
1974. rug
1975. misery
1976. awkward
1977. rival
1978. roast
1979. deed
1980. preference
1981. explosion
1982. theatrical
1983. cultivate
1984. collector
1985. miserable
1986. wrist
1987. rabbit
1988. accustom
1989. tide
1990. insult
1991. thumb
1992. lump
1993. annoy
1994. toy
1995. heal
1996. shallow
1997. repetition
1998. soup
1999. whistle
2000. scenery
2001. apple
2002. offense
2003. cork
2004. ripe
2005. temper
2006. sore
2007. pinch
2008. diamond
2009. razor
2010. imaginative
2011. hook
2012. copper
2013. landlord
2014. influential
2015. rot
2016. hollow
2017. enclose
2018. harden
2019. wicked
2020. stiffen
2021. silk
2022. upright
2023. selfish
2024. stripe
2025. pig
2026. inward
2027. excellence
2028. rake
2029. purple
2030. hasten
2031. shorten
2032. applause
2033. ache
2034. apology
2035. knot
2036. nephew
2037. cushion
2038. drown
2039. nursery
2040. pint
2041. fierce
2042. imitate
2043. aloud
2044. gaiety
2045. robbery
2046. tighten
2047. perfection
2048. scorn
2049. whoever
2050. trunk
2051. wool
2052. sailor
2053. competitor
2054. moonlight
2055. deer
2056. bean
2057. everyday
2058. drawer
2059. disregard
2060. nowadays
2061. patriotic
2062. tin
2063. penny
2064. cage
2065. pardon
2066. lately
2067. offend
2068. coarse
2069. spoil
2070. horizontal
2071. sting
2072. ditch
2073. librarian
2074. meantime
2075. cough
2076. deaf
2077. sword
2078. messenger
2079. vain
2080. castle
2081. elastic
2082. comb
2083. rod
2084. widen
2085. sorrow
2086. inventor
2087. cliff
2088. umbrella
2089. interruption
2090. merry
2091. gallon
2092. conquest
2093. headache
2094. tailor
2095. bucket
2096. scent

2097. signature
2098. cart
2099. darken
2100. sometime
2101. applaud
2102. underneath
2103. hello
2104. pretense
2105. descent
2106. conquer
2107. framework
2108. confidential
2109. adoption
2110. disgust
2111. waist
2112. momentary
2113. receipt
2114. pearl
2115. ray
2116. lazy
2117. limb
2118. grammatical
2119. beast
2120. monkey
2121. jewel
2122. persuasion
2123. obedience
2124. sock
2125. vowel
2126. hammer
2127. inn
2128. chimney
2129. dissatisfaction
2130. annoyance
2131. ornament
2132. honesty
2133. outward
2134. sharpen
2135. handkerchief
2136. greed
2137. heavenly
2138. thirst
2139. niece
2140. spill
2141. loaf
2142. wheat
2143. worm
2144. secrecy
2145. rude
2146. heighten
2147. flatten
2148. loosen
2149. cheese
2150. rivalry
2151. royalty
2152. discontent
2153. complication
2154. fright
2155. indoor
2156. flour
2157. actress
2158. congratulation
2159. ounce
2160. fry
2161. everlasting
2162. goat
2163. ink
2164. disappearance
2165. reproduction
2166. thicken
2167. avoidance
2168. spoon
2169. strap
2170. deceive
2171. lengthen
2172. revenge
2173. correction
2174. descendant
2175. hesitation
2176. spade
2177. basin
2178. weed
2179. omission
2180. old-fashioned
2181. bicycle
2182. breadth
2183. photography
2184. coward
2185. mat
2186. rejoice
2187. cheat
2188. congratulate
2189. discomfort
2190. enclosure
2191. attentive
2192. paw
2193. overflow
2194. dissatisfy
2195. multiplication
2196. whichever
2197. tidy
2198. bribe
2199. mend
2200. stocking
2201. feast
2202. nuisance
2203. thorn
2204. tame
2205. inclusive
2206. homemade
2207. handwriting
2208. chalk
2209. sour
2210. slippery
2211. procession
2212. ripen
2213. jealous
2214. jealousy
2215. liar
2216. homecoming
2217. barber
2218. whiten
2219. berry
2220. lighten
2221. pigeon
2222. hinder
2223. bravery
2224. baggage
2225. noun
2226. amongst
2227. grammar
2228. cultivation
2229. companionship
2230. rubbish
2231. modesty
2232. woolen
2233. deepen
2234. pastry
2235. cupboard
2236. quart
2237. canal
2238. notebook
2239. deceit
2240. parcel
2241. brighten
2242. moderation
2243. punctual
2244. hurrah
2245. lipstick
2246. uppermost
2247. fatten
2248. conqueror
2249. hindrance
2250. cowardice
2251. obedient
2252. saucer
2253. madden
2254. scold
2255. weekday
2256. rotten
2257. disrespect
2258. widower
2259. deafen
2260. donkey
2261. businesslike
2262. motherhood
2263. sadden
2264. handshake
2265. calculator
2266. headdress
2267. scissors
2268. translator
2269. possessor
2270. shilling
2271. redden
2272. motherly
2273. whose
2274. cultivator
2275. whom
2276. homework
2277. electrician
2278. oar
2279. bribery
2280. sweeten
2281. sow
2282. pronunciation
2283. beak
2284. plural

APPENDIX B

The Academic Word List

(Coxhead, 2000)

This list contains 570 head words of the families in the Academic Word List. The word list has been divided into sublists based on the frequency of occurrence of the words in the Academic Corpus. The words in Sublist 1 occur more frequently in the corpus than the other words in the list. Sublist 2 occurs with the next highest frequency.

Sublist 1

analyse approach area assess assume authority available benefit concept consist constitute context contract create data define derive distribute economy environment establish estimate evident export factor finance formula function identify income indicate individual interpret involve issue labour legal legislate major method occur percent period policy principle proceed process require research respond role section sector significant similar source specific structure theory vary

Sublist 2

achieve acquire administrate affect appropriate aspect assist category chapter commission community complex compute conclude conduct consequent construct consume credit culture design distinct element equate evaluate feature final focus impact injure institute invest item journal maintain normal obtain participate perceive positive potential previous primary purchase range region regulate relevant reside resource restrict secure seek select site strategy survey text tradition transfer

Sublist 3

alternative circumstance comment compensate component consent considerable constant constrain contribute convene coordinate core corporate correspond criteria deduce demonstrate document dominate emphasis ensure exclude framework fund illustrate immigrate imply initial instance

interact justify layer link locate maximise minor negate outcome partner philosophy physical proportion publish react register rely remove scheme sequence sex shift specify sufficient task technical technique technology valid volume

Sublist 4

access adequate annual apparent approximate attitude attribute civil code commit communicate concentrate confer contrast cycle debate despite dimension domestic emerge error ethnic goal grant hence hypothesis implement implicate impose integrate internal investigate job label mechanism obvious occupy option output overall parallel parameter phase predict principal prior professional project promote regime resolve retain series statistic status stress subsequent sum summary undertake

Sublist 5

academy adjust alter amend aware capacity challenge clause compound conflict consult contact decline discrete draft enable energy enforce entity equivalent evolve expand expose external facilitate fundamental generate generation image liberal licence logic margin medical mental modify monitor network notion objective orient perspective precise prime psychology pursue ratio reject revenue stable style substitute sustain symbol target transit trend version welfare whereas

Sublist 6

abstract accurate acknowledge aggregate allocate assign attach author bond brief capable cite cooperate discriminate display diverse domain edit enhance estate exceed expert explicit federal fee flexible furthermore gender ignorant incentive incidence incorporate index inhibit initiate input instruct intelligence interval lecture migrate minimum ministry motive neutral nevertheless overseas precede presume rational recover reveal scope subsidy tape trace transform transport underlie utilise

Sublist 7

adapt adult advocate aid channel chemical classic comprehensive comprise confirm contrary convert couple decade definite deny differentiate dispose dynamic eliminate empirical equip extract file finite foundation globe grade guarantee hierarchy identical ideology infer innovate insert intervene isolate media mode paradigm phenomenon priority prohibit publication quote release reverse simulate sole somewhat submit successor survive thesis topic transmit ultimate unique visible voluntary

Sublist 8

abandon accompany accumulate ambiguous append appreciate arbitrary automate bias chart clarify commodity complement conform contemporary contradict crucial currency denote detect deviate displace drama eventual exhibit exploit fluctuate guideline highlight implicit induce inevitable infrastructure inspect intense manipulate minimise nuclear offset paragraph plus practitioner predominant prospect radical random reinforce restore revise schedule tense terminate theme thereby uniform vehicle via virtual visual widespread

Sublist 9

accommodate analogy anticipate assure attain behalf bulk cease coherent coincide commence compatible concurrent confine controversy converse device devote diminish distort duration erode ethic format founded inherent insight integral intermediate manual mature mediate medium military minimal mutual norm overlap passive portion preliminary protocol qualitative refine relax restrain revolution rigid route scenario sphere subordinate supplement suspend team temporary trigger unify violate vision

Sublist 10

adjacent albeit assemble collapse colleague compile conceive convince depress encounter enormous forthcoming incline integrity intrinsic invoke levy likewise nonetheless notwithstanding odd ongoing panel persist pose reluctance so-called straightforward undergo whereby

APPENDIX C

The Word Frequency List of the Pharmaceutical Corpus (Raw data)

Rank	Word	Freq.	%	Rank	Word	Freq.	%
1.	the	12,699	3.68	2.	of	12,530	3.63
3.	in	9,167	2.66	4.	and	8,572	2.49
5.	to	6,016	1.74	6.	with	5,053	1.47
7.	be	4,511	1.31	8.	is	4,259	1.23
9.	or	4,116	1.19	10.	a	4,055	1.18
11.	patients	3,250	0.94	12.	mg	2,685	0.78
13.	should	2,616	0.76	14.	for	2,524	0.73
15.	may	2,059	0.6	16.	dose	1,894	0.55
17.	not	1,860	0.54	18.	treatment	1,656	0.48
19.	as	1,642	0.48	20.	been	1,581	0.46
21.	by	1,535	0.45	22.	are	1,503	0.44
23.	have	1,371	0.4	24.	at	1,283	0.37
25.	on	1,209	0.35	26.	therapy	1,160	0.34
27.	administration	1,076	0.31	28.	use	1,061	0.31
29.	dosage	1,037	0.3	30.	has	1,030	0.3
31.	was	1,003	0.29	32.	after	983	0.28
33.	other	966	0.28	34.	that	950	0.28
35.	no	942	0.27	36.	ml	936	0.27
37.	daily	924	0.27	38.	drug	919	0.27
39.	reactions	902	0.26	40.	if	899	0.26
41.	an	893	0.26	42.	effects	831	0.24
43.	doses	830	0.24	44.	were	826	0.24
45.	from	817	0.24	46.	adverse	790	0.23
47.	blood	786	0.23	48.	renal	779	0.23
49.	it	736	0.21	50.	during	733	0.21
51.	when	717	0.21	52.	clinical	689	0.2
53.	following	683	0.2	54.	studies	666	0.19
55.	x	658	0.19	56.	reported	657	0.19
57.	day	655	0.19	58.	recommended	655	0.19
59.	this	655	0.19	60.	administered	654	0.19
61.	used	654	0.19	62.	severe	636	0.18
63.	effect	622	0.18	64.	plasma	620	0.18
65.	than	618	0.18	66.	see	615	0.18
67.	these	613	0.18	68.	increased	572	0.17
69.	oral	567	0.16	70.	kg	549	0.16
71.	disease	526	0.15	72.	there	524	0.15
73.	which	521	0.15	74.	eg	513	0.15
75.	hrs	513	0.15	76.	occur	513	0.15
77.	observed	508	0.15	78.	symptoms	505	0.15
79.	treated	505	0.15	80.	function	501	0.15
81.	pregnancy	496	0.14	82.	mcg	484	0.14
83.	including	477	0.14	84.	tab	472	0.14
85.	given	464	0.13	86.	can	460	0.13
87.	also	455	0.13	88.	days	453	0.13
89.	iv	453	0.13	90.	precautions	449	0.13
91.	children	441	0.13	92.	drugs	433	0.13
93.	infusion	430	0.12	94.	patient	420	0.12
95.	concentrations	419	0.12	96.	injection	416	0.12
97.	levels	410	0.12	98.	acute	408	0.12
99.	any	408	0.12	100.	cases	406	0.12

Rank	Word	Freq.	%	Rank	Word	Freq.	%
101.	who	406	0.12	102.	pressure	402	0.12
103.	packing	396	0.11	104.	hepatic	395	0.11
105.	serum	393	0.11	106.	skin	387	0.11
107.	liver	369	0.11	108.	times	364	0.11
109.	only	359	0.1	110.	increase	358	0.1
111.	levofloxacin	354	0.1	112.	solution	344	0.1
113.	more	338	0.1	114.	indications	337	0.1
115.	due	334	0.1	116.	risk	329	0.1
117.	single	323	0.09	118.	but	319	0.09
119.	impairment	318	0.09	120.	must	317	0.09
121.	pain	317	0.09	122.	time	310	0.09
123.	c	309	0.09	124.	once	308	0.09
125.	min	306	0.09	126.	approximately	305	0.09
127.	body	303	0.09	128.	under	302	0.09
129.	all	300	0.09	130.	heart	296	0.09
131.	interaction	296	0.09	132.	its	295	0.09
133.	hypersensitivity	294	0.09	134.	weeks	292	0.08
135.	years	292	0.08	136.	associated	288	0.08
137.	response	288	0.08	138.	usually	288	0.08
139.	presentation	285	0.08	140.	cardiac	284	0.08
141.	within	284	0.08	142.	rate	282	0.08
143.	study	282	0.08	144.	infections	279	0.08
145.	those	279	0.08	146.	up	277	0.08
147.	months	276	0.08	148.	manufacturer	274	0.08
149.	such	274	0.08	150.	toxicity	274	0.08
151.	distributor	272	0.08	152.	receiving	271	0.08
153.	cancer	269	0.08	154.	taken	268	0.08
155.	however	267	0.08	156.	life	267	0.08
157.	therefore	267	0.08	158.	some	266	0.08
159.	before	265	0.08	160.	agents	261	0.08
161.	high	261	0.08	162.	caution	260	0.08
163.	each	260	0.08	164.	failure	259	0.08
165.	combination	257	0.07	166.	contraindication	257	0.07
167.	adults	256	0.07	168.	disorders	256	0.07
169.	chronic	250	0.07	170.	most	250	0.07
171.	normal	250	0.07	172.	reduced	250	0.07
173.	related	250	0.07	174.	weight	250	0.07
175.	women	249	0.07	176.	pharmacokinetics	247	0.07
177.	reduction	247	0.07	178.	about	245	0.07
179.	hr	245	0.07	180.	potential	245	0.07
181.	activity	243	0.07	182.	rarely	243	0.07
183.	significant	243	0.07	184.	overdosage	242	0.07
185.	elderly	241	0.07	186.	greater	241	0.07
187.	gastrointestinal	238	0.07	188.	human	237	0.07
189.	known	237	0.07	190.	tablet	237	0.07
191.	breast	236	0.07	192.	into	235	0.07
193.	necessary	235	0.07	194.	tablets	234	0.07
195.	equal	233	0.07	196.	events	232	0.07
197.	data	231	0.07	198.	without	231	0.07
199.	actions	230	0.07	200.	shown	230	0.07
201.	absorption	229	0.07	202.	systemic	229	0.07
203.	twice	229	0.07	204.	hypertension	228	0.07
205.	sodium	228	0.07	206.	well	228	0.07
207.	active	226	0.07	208.	nausea	226	0.07
209.	rats	225	0.07	210.	clearance	223	0.06

Rank	Word	Freq.	%	Rank	Word	Freq.	%
211.	trials	222	0.06	212.	taxol	218	0.06
213.	system	217	0.06	214.	l	216	0.06
215.	action	215	0.06	216.	first	215	0.06
217.	over	214	0.06	218.	cefaclor	212	0.06
219.	half	210	0.06	220.	between	208	0.06
221.	controlled	208	0.06	222.	considered	204	0.06
223.	since	204	0.06	224.	concentration	203	0.06
225.	impaired	203	0.06	226.	safety	200	0.06
227.	creatinine	199	0.06	228.	lactation	199	0.06
229.	asthma	198	0.06	230.	possible	198	0.06
231.	rare	198	0.06	232.	required	198	0.06
233.	table	198	0.06	234.	appropriate	197	0.06
235.	higher	197	0.06	236.	therapeutic	197	0.06
237.	form	196	0.06	238.	long	194	0.06
239.	both	191	0.06	240.	concomitant	191	0.06
241.	mild	191	0.06	242.	monitored	191	0.06
243.	because	190	0.06	244.	class	190	0.06
245.	side	190	0.06	246.	cell	189	0.05
247.	contains	189	0.05	248.	g	189	0.05
249.	maximum	189	0.05	250.	nasal	189	0.05
251.	control	188	0.05	252.	corticosteroids	188	0.05
253.	hypotension	188	0.05	254.	range	188	0.05
255.	vomiting	188	0.05	256.	incidence	187	0.05
257.	using	187	0.05	258.	discontinued	186	0.05
259.	elimination	185	0.05	260.	allergic	184	0.05
261.	every	184	0.05	262.	mean	182	0.05
263.	conditions	177	0.05	264.	water	177	0.05
265.	acid	176	0.05	266.	pregnant	176	0.05
267.	tract	176	0.05	268.	result	175	0.05
269.	cause	174	0.05	270.	prolonged	174	0.05
271.	one	173	0.05	272.	changes	172	0.05
273.	lower	170	0.05	274.	w	170	0.05
275.	excreted	169	0.05	276.	had	169	0.05
277.	adult	168	0.05	278.	bleeding	168	0.05
279.	low	168	0.05	280.	until	168	0.05
281.	clinically	167	0.05	282.	distaclor	167	0.05
283.	metabolism	167	0.05	284.	headache	166	0.05
285.	phase	166	0.05	286.	urine	166	0.05
287.	dizziness	165	0.05	288.	syndrome	165	0.05
289.	vitro	165	0.05	290.	established	164	0.05
291.	p	164	0.05	292.	does	163	0.05
293.	signs	163	0.05	294.	test	162	0.05
295.	urinary	162	0.05	296.	subjects	161	0.05
297.	effective	160	0.05	298.	evidence	160	0.05
299.	decreased	159	0.05	300.	either	159	0.05
301.	indicated	159	0.05	302.	contents	157	0.05
303.	do	157	0.05	304.	interactions	157	0.05
305.	experience	156	0.05	306.	generally	156	0.05
307.	occurred	156	0.05	308.	although	155	0.04
309.	containing	155	0.04	310.	duration	155	0.04
311.	history	155	0.04	312.	metabolites	155	0.04
313.	prior	155	0.04	314.	vial	155	0.04
315.	exposure	154	0.04	316.	fluid	154	0.04
317.	infection	154	0.04	318.	peripheral	154	0.04
319.	group	153	0.04	320.	peak	153	0.04

Rank	Word	Freq.	%	Rank	Word	Freq.	%
321.	very	153	0.04	322.	least	152	0.04
323.	results	152	0.04	324.	term	152	0.04
325.	total	152	0.04	326.	discontinuation	151	0.04
327.	particularly	151	0.04	328.	week	151	0.04
329.	less	150	0.04	330.	inflammatory	149	0.04
331.	initial	149	0.04	332.	moderate	149	0.04
333.	similar	149	0.04	334.	disturbances	148	0.04
335.	laboratory	148	0.04	336.	milk	147	0.04
337.	adalat	144	0.04	338.	respiratory	144	0.04
339.	rash	143	0.04	340.	reaction	143	0.04
341.	s	143	0.04	342.	decrease	142	0.04
343.	respectively	142	0.04	344.	tests	142	0.04
345.	monitoring	141	0.04	346.	serious	141	0.04
347.	state	141	0.04	348.	occurs	140	0.04
349.	received	140	0.04	350.	salbutamol	140	0.04
351.	cells	139	0.04	352.	mims	139	0.04
353.	they	139	0.04	354.	caused	138	0.04
355.	common	138	0.04	356.	fever	138	0.04
357.	hours	138	0.04	358.	period	138	0.04
359.	while	138	0.04	360.	myocardial	137	0.04
361.	orally	137	0.04	362.	white	137	0.04
363.	available	136	0.04	364.	mr	136	0.04
365.	store	136	0.04	366.	general	135	0.04
367.	muscle	135	0.04	368.	nifedipine	135	0.04
369.	potassium	135	0.04	370.	against	134	0.04
371.	based	134	0.04	372.	maintenance	134	0.04
373.	reports	134	0.04	374.	reversible	134	0.04
375.	strains	134	0.04	376.	clarithromycin	133	0.04
377.	corticosteroid	133	0.04	378.	storage	133	0.04
379.	placebo	132	0.04	380.	will	132	0.04
381.	like	131	0.04	382.	ability	130	0.04
383.	efficacy	130	0.04	384.	bone	129	0.04
385.	casodex	129	0.04	386.	tachycardia	129	0.04
387.	values	129	0.04	388.	annual	128	0.04
389.	inhalation	127	0.04	390.	agent	126	0.04
391.	dosing	126	0.04	392.	performed	126	0.04
393.	multidose	125	0.04	394.	positive	125	0.04
395.	compared	124	0.04	396.	dry	124	0.04
397.	inhaled	124	0.04	398.	per	124	0.04
399.	require	124	0.04	400.	sandimmun	124	0.04
401.	taking	124	0.04	402.	pen	123	0.04
403.	photo	123	0.04	404.	increases	122	0.04
405.	machinery	122	0.04	406.	price	122	0.04
407.	excretion	121	0.04	408.	immediately	121	0.04
409.	hcl	120	0.03	410.	insufficiency	120	0.03
411.	intron	120	0.03	412.	seen	120	0.03
413.	usual	120	0.03	414.	divided	119	0.03
415.	level	119	0.03	416.	repeated	119	0.03
417.	depression	118	0.03	418.	neutropenia	118	0.03
419.	patient's	118	0.03	420.	spray	118	0.03
421.	case	117	0.03	422.	co	117	0.03
423.	especially	117	0.03	424.	followed	117	0.03
425.	occasionally	117	0.03	426.	small	117	0.03
427.	anti	116	0.03	428.	digoxin	116	0.03
429.	drive	116	0.03	430.	induced	116	0.03

Rank	Word	Freq.	%	Rank	Word	Freq.	%
431.	post	116	0.03	432.	specific	116	0.03
433.	inhibition	115	0.03	434.	measures	115	0.03
435.	special	115	0.03	436.	angina	114	0.03
437.	d	114	0.03	438.	diarrhea	114	0.03
439.	few	114	0.03	440.	transient	114	0.03
441.	animal	113	0.03	442.	cap	113	0.03
443.	their	113	0.03	444.	volume	113	0.03
445.	local	111	0.03	446.	resistant	111	0.03
447.	thrombocytopenia	111	0.03	448.	coronary	110	0.03
449.	hepatitis	110	0.03	450.	pharmacology	110	0.03
451.	physician	110	0.03	452.	xeloda	110	0.03
453.	according	109	0.03	454.	auc	109	0.03
455.	description	109	0.03	456.	loss	109	0.03
457.	previously	109	0.03	458.	same	109	0.03
459.	through	109	0.03	460.	tissue	109	0.03
461.	i	108	0.03	462.	mouth	108	0.03
463.	cisplatin	107	0.03	464.	nsaids	107	0.03
465.	release	107	0.03	466.	grade	106	0.03
467.	operate	106	0.03	468.	relief	106	0.03
469.	tsp	106	0.03	470.	alone	105	0.03
471.	antibiotics	105	0.03	472.	further	105	0.03
473.	several	105	0.03	474.	cardiovascular	104	0.03
475.	healthy	104	0.03	476.	medical	104	0.03
477.	nursing	104	0.03	478.	adequate	103	0.03
479.	cough	103	0.03	480.	count	103	0.03
481.	gi	103	0.03	482.	metabolite	103	0.03
483.	coated	102	0.03	484.	inhibitors	102	0.03
485.	lung	102	0.03	486.	severity	102	0.03
487.	significantly	101	0.03	488.	susceptible	101	0.03
489.	where	101	0.03	490.	below	100	0.03
491.	childn	100	0.03	492.	demonstrated	100	0.03
493.	mice	100	0.03	494.	frequent	99	0.03
495.	steady	99	0.03	496.	adjustment	98	0.03
497.	event	98	0.03	498.	neoral	98	0.03
499.	t	98	0.03	500.	diuretics	97	0.03
501.	expected	97	0.03	502.	frequently	97	0.03
503.	im	97	0.03	504.	non	97	0.03
505.	carcinoma	96	0.03	506.	distribution	96	0.03
507.	edema	96	0.03	508.	line	96	0.03
509.	subsequent	96	0.03	510.	gastric	95	0.03
511.	toxic	95	0.03	512.	ventolin	95	0.03
513.	glucose	94	0.03	514.	herceptin	94	0.03
515.	pneumoniae	94	0.03	516.	properties	94	0.03
517.	protein	94	0.03	518.	retention	94	0.03
519.	rhinitis	94	0.03	520.	symptomatic	94	0.03
521.	ventricular	94	0.03	522.	absorbed	93	0.03
523.	care	93	0.03	524.	carefully	93	0.03
525.	concomitantly	93	0.03	526.	continued	93	0.03
527.	flow	93	0.03	528.	frequency	93	0.03
529.	include	93	0.03	530.	susceptibility	93	0.03
531.	achieved	92	0.03	532.	age	92	0.03
533.	arthritis	92	0.03	534.	platelet	92	0.03
535.	standard	92	0.03	536.	upon	92	0.03
537.	antihypertensive	91	0.03	538.	benefit	91	0.03
539.	caelyx	91	0.03	540.	calcium	91	0.03

Rank	Word	Freq.	%	Rank	Word	Freq.	%
541.	enzyme	91	0.03	542.	pulmonary	91	0.03
543.	rapid	91	0.03	544.	site	91	0.03
545.	streptococcus	91	0.03	546.	abnormal	90	0.03
547.	abnormalities	90	0.03	548.	addition	90	0.03
549.	b	90	0.03	550.	chemotherapy	90	0.03
551.	kidney	90	0.03	552.	marrow	90	0.03
553.	suspension	90	0.03	554.	then	90	0.03
555.	ace	89	0.03	556.	found	89	0.03
557.	median	89	0.03	558.	medication	89	0.03
559.	yr	89	0.03	560.	cns	88	0.03
561.	dysfunction	88	0.03	562.	initiated	88	0.03
563.	loniten	88	0.03	564.	propionate	88	0.03
565.	pruritus	88	0.03	566.	warnings	88	0.03
567.	avoid	87	0.03	568.	being	87	0.03
569.	colitis	87	0.03	570.	fluticasone	87	0.03
571.	n	87	0.03	572.	urticaria	87	0.03
573.	binding	86	0.02	574.	h	86	0.02
575.	hydrochloride	86	0.02	576.	improvement	86	0.02
577.	intervals	86	0.02	578.	light	86	0.02
579.	obtained	86	0.02	580.	sc	86	0.02
581.	sr	86	0.02	582.	your	86	0.02
583.	abdominal	85	0.02	584.	animals	85	0.02
585.	dependent	85	0.02	586.	gabapentin	85	0.02
587.	out	85	0.02	588.	you	85	0.02
589.	cefprozil	84	0.02	590.	did	84	0.02
591.	ng	84	0.02	592.	unchanged	84	0.02
593.	diclofenac	83	0.02	594.	food	83	0.02
595.	m	83	0.02	596.	adjusted	82	0.02
597.	dobutamine	82	0.02	598.	excessive	82	0.02
599.	management	82	0.02	600.	pharmacokinetic	82	0.02
601.	careful	81	0.02	602.	chloride	81	0.02
603.	diuretic	81	0.02	604.	equivalent	81	0.02
605.	humans	81	0.02	606.	isolated	81	0.02
607.	overdose	81	0.02	608.	rapidly	81	0.02
609.	starting	81	0.02	610.	thus	81	0.02
611.	bioavailability	80	0.02	612.	cold	80	0.02
613.	dilution	80	0.02	614.	pediatric	80	0.02
615.	present	80	0.02	616.	syr	80	0.02
617.	via	80	0.02	618.	adrenal	79	0.02
619.	extent	79	0.02	620.	infarction	79	0.02
621.	presence	79	0.02	622.	surgery	79	0.02
623.	testing	79	0.02	624.	topical	79	0.02
625.	voltaren	79	0.02	626.	contraindicated	78	0.02
627.	glaucoma	78	0.02	628.	parenteral	78	0.02
629.	regimen	78	0.02	630.	resulting	78	0.02
631.	retard	78	0.02	632.	temperature	78	0.02
633.	unless	78	0.02	634.	valdecoxib	78	0.02
635.	condition	77	0.02	636.	likely	77	0.02
637.	organisms	77	0.02	638.	powder	77	0.02
639.	survival	77	0.02	640.	two	77	0.02
641.	vascular	77	0.02	642.	amp	76	0.02
643.	dermatitis	76	0.02	644.	information	76	0.02
645.	mainly	76	0.02	646.	mm	76	0.02
647.	multiple	76	0.02	648.	negative	76	0.02
649.	zuellig	76	0.02	650.	advanced	75	0.02

Rank	Word	Freq.	%	Rank	Word	Freq.	%
651.	influenzae	75	0.02	652.	major	75	0.02
653.	month	75	0.02	654.	particular	75	0.02
655.	reduce	75	0.02	656.	solutions	75	0.02
657.	whether	75	0.02	658.	baseline	74	0.02
659.	cream	74	0.02	660.	exercised	74	0.02
661.	haemophilus	74	0.02	662.	individual	74	0.02
663.	infants	74	0.02	664.	inhibitor	74	0.02
665.	oedema	74	0.02	666.	previous	74	0.02
667.	status	74	0.02	668.	vision	74	0.02
669.	acetate	73	0.02	670.	bronchospasm	73	0.02
671.	course	73	0.02	672.	develop	73	0.02
673.	diabetes	73	0.02	674.	diluted	73	0.02
675.	e	73	0.02	676.	even	73	0.02
677.	growth	73	0.02	678.	inhaler	73	0.02
679.	interferon	73	0.02	680.	nervous	73	0.02
681.	overall	73	0.02	682.	paclitaxel	73	0.02
683.	volunteers	73	0.02	684.	area	72	0.02
685.	could	72	0.02	686.	development	72	0.02
687.	dogs	72	0.02	688.	nizoral	72	0.02
689.	procedures	72	0.02	690.	receptor	72	0.02
691.	salmeterol	72	0.02	692.	short	72	0.02
693.	warfarin	72	0.02	694.	alcohol	71	0.02
695.	conducted	71	0.02	696.	depending	71	0.02
697.	diarrhoea	71	0.02	698.	indicate	71	0.02
699.	intestinal	71	0.02	700.	limited	71	0.02
701.	obstruction	71	0.02	702.	tolerated	71	0.02
703.	upper	71	0.02	704.	disorder	70	0.02
705.	doxorubicin	70	0.02	706.	onset	70	0.02
707.	pneumonia	70	0.02	708.	stable	70	0.02
709.	start	70	0.02	710.	budesonide	69	0.02
711.	central	69	0.02	712.	cr	69	0.02
713.	gram	69	0.02	714.	important	69	0.02
715.	number	69	0.02	716.	ovarian	69	0.02
717.	parameters	69	0.02	718.	paraplatin	69	0.02
719.	resistance	69	0.02	720.	seizures	69	0.02
721.	teratogenic	69	0.02	722.	tissues	69	0.02
723.	additional	68	0.02	724.	anemia	68	0.02
725.	discontinue	68	0.02	726.	early	68	0.02
727.	platinol	68	0.02	728.	producing	68	0.02
729.	progression	68	0.02	730.	relationship	68	0.02
731.	syrup	68	0.02	732.	avoided	67	0.02
733.	bronchial	67	0.02	734.	change	67	0.02
735.	doctor	67	0.02	736.	large	67	0.02
737.	made	67	0.02	738.	metastatic	67	0.02
739.	r	67	0.02	740.	suppression	67	0.02
741.	type	67	0.02	742.	appear	66	0.02
743.	cerebral	66	0.02	744.	damage	66	0.02
745.	dextrose	66	0.02	746.	film	66	0.02
747.	ph	66	0.02	748.	primary	66	0.02
749.	relevant	66	0.02	750.	anticoagulants	65	0.02
751.	bextra	65	0.02	752.	dialysis	65	0.02
753.	elevated	65	0.02	754.	elidel	65	0.02
755.	evaluated	65	0.02	756.	give	65	0.02
757.	inhibit	65	0.02	758.	pectoris	65	0.02
759.	prevention	65	0.02	760.	smaller	65	0.02

Rank	Word	Freq.	%	Rank	Word	Freq.	%
761.	sulfate	65	0.02	762.	withdrawal	65	0.02
763.	certain	64	0.02	764.	closely	64	0.02
765.	fertility	64	0.02	766.	impair	64	0.02
767.	male	64	0.02	768.	many	64	0.02
769.	potent	64	0.02	770.	prevent	64	0.02
771.	studied	64	0.02	772.	tid	64	0.02
773.	amount	63	0.02	774.	application	63	0.02
775.	bradycardia	63	0.02	776.	constipation	63	0.02
777.	counts	63	0.02	778.	described	63	0.02
779.	experienced	63	0.02	780.	maleate	63	0.02
781.	meloxicam	63	0.02	782.	noted	63	0.02
783.	possibility	63	0.02	784.	preclinical	63	0.02
785.	rabbits	63	0.02	786.	room	63	0.02
787.	combined	62	0.02	788.	complete	62	0.02
789.	erythema	62	0.02	790.	produce	62	0.02
791.	threatening	62	0.02	792.	blocking	61	0.02
793.	concurrent	61	0.02	794.	included	61	0.02
795.	leukopenia	61	0.02	796.	mothers	61	0.02
797.	need	61	0.02	798.	note	61	0.02
799.	primarily	61	0.02	800.	tumors	61	0.02
801.	warning	61	0.02	802.	determined	60	0.02
803.	direct	60	0.02	804.	diseases	60	0.02
805.	follows	60	0.02	806.	irritation	60	0.02
807.	metabolic	60	0.02	808.	nolvadex	60	0.02
809.	pre	60	0.02	810.	rheumatoid	60	0.02
811.	slight	60	0.02	812.	slightly	60	0.02
813.	trial	60	0.02	814.	ulceration	60	0.02
815.	advised	59	0.02	816.	anorexia	59	0.02
817.	arrhythmias	59	0.02	818.	different	59	0.02
819.	doxazosin	59	0.02	820.	drowsiness	59	0.02
821.	highest	59	0.02	822.	instructions	59	0.02
823.	intravenous	59	0.02	824.	longer	59	0.02
825.	miu	59	0.02	826.	regular	59	0.02
827.	take	59	0.02	828.	vivo	59	0.02
829.	alopecia	58	0.02	830.	angioedema	58	0.02
831.	average	58	0.02	832.	blockers	58	0.02
833.	dyspnea	58	0.02	834.	inhibits	58	0.02
835.	intranasal	58	0.02	836.	magnesium	58	0.02
837.	needed	58	0.02	838.	suitable	58	0.02
839.	toxicology	58	0.02	840.	above	57	0.02
841.	bronchitis	57	0.02	842.	compound	57	0.02
843.	inactive	57	0.02	844.	measured	57	0.02
845.	quinolones	57	0.02	846.	sarcoma	57	0.02
847.	secretion	57	0.02	848.	showed	57	0.02
849.	sp	57	0.02	850.	affected	56	0.02
851.	antibiotic	56	0.02	852.	apparent	56	0.02
853.	chest	56	0.02	854.	cravit	56	0.02
855.	diprospan	56	0.02	856.	disk	56	0.02
857.	exceed	56	0.02	858.	heparin	56	0.02
859.	initially	56	0.02	860.	isotretinoin	56	0.02
861.	mechanism	56	0.02	862.	mic	56	0.02
863.	oxaliplatin	56	0.02	864.	oxygen	56	0.02
865.	prophylaxis	56	0.02	866.	soluble	56	0.02
867.	ulcer	56	0.02	868.	usage	56	0.02
869.	aids	55	0.02	870.	lead	55	0.02

Rank	Word	Freq.	%	Rank	Word	Freq.	%
871.	marked	55	0.02	872.	occurring	55	0.02
873.	order	55	0.02	874.	prescribed	55	0.02
875.	rates	55	0.02	876.	resulted	55	0.02
877.	year	55	0.02	878.	among	54	0.02
879.	brain	54	0.02	880.	cannot	54	0.02
881.	cimetidine	54	0.02	882.	fenamon	54	0.02
883.	fetus	54	0.02	884.	findings	54	0.02
885.	ie	54	0.02	886.	lesions	54	0.02
887.	maintained	54	0.02	888.	population	54	0.02
889.	potentially	54	0.02	890.	protect	54	0.02
891.	shock	54	0.02	892.	started	54	0.02
893.	whole	54	0.02	894.	close	53	0.02
895.	contraindications	53	0.02	896.	cox	53	0.02
897.	detected	53	0.02	898.	diltiazem	53	0.02
899.	docetaxel	53	0.02	900.	inhibitory	53	0.02
901.	penicillin	53	0.02	902.	platinum	53	0.02
903.	profile	53	0.02	904.	reached	53	0.02
905.	reproductive	53	0.02	906.	staphylococcus	53	0.02
907.	synthesis	53	0.02	908.	adrenergic	52	0.02
909.	advisable	52	0.02	910.	affect	52	0.02
911.	antacids	52	0.02	912.	congestive	52	0.02
913.	cytochrome	52	0.02	914.	decreases	52	0.02
915.	lamivudine	52	0.02	916.	marketing	52	0.02
917.	meals	52	0.02	918.	mometasone	52	0.02
919.	preparation	52	0.02	920.	supervision	52	0.02
921.	tumor	52	0.02	922.	angiotensin	51	0.01
923.	bound	51	0.01	924.	continuous	51	0.01
925.	dna	51	0.01	926.	essential	51	0.01
927.	formation	51	0.01	928.	gradually	51	0.01
929.	herbesser	51	0.01	930.	inhibited	51	0.01
931.	sensitive	51	0.01	932.	slow	51	0.01
933.	sometimes	51	0.01	934.	supportive	51	0.01
935.	surface	51	0.01	936.	v	51	0.01
937.	visual	51	0.01	938.	blurred	50	0.01
939.	campto	50	0.01	940.	courses	50	0.01
941.	face	50	0.01	942.	fetal	50	0.01
943.	full	50	0.01	944.	iii	50	0.01
945.	inflammation	50	0.01	946.	ingredients	50	0.01
947.	methotrexate	50	0.01	948.	rat	50	0.01
949.	route	50	0.01	950.	taste	50	0.01
951.	topotecan	50	0.01	952.	tumour	50	0.01
953.	accumulation	49	0.01	954.	al	49	0.01
955.	bilirubin	49	0.01	956.	block	49	0.01
957.	cephalosporins	49	0.01	958.	circulatory	49	0.01
959.	enzymes	49	0.01	960.	f	49	0.01
961.	free	49	0.01	962.	glivec	49	0.01
963.	groups	49	0.01	964.	infrequently	49	0.01
965.	often	49	0.01	966.	report	49	0.01
967.	various	49	0.01	968.	absence	48	0.01
969.	acting	48	0.01	970.	bacterial	48	0.01
971.	date	48	0.01	972.	expectorant	48	0.01
973.	factor	48	0.01	974.	furoate	48	0.01
975.	hair	48	0.01	976.	indicates	48	0.01
977.	preparations	48	0.01	978.	produced	48	0.01
979.	seretide	48	0.01	980.	so	48	0.01

Rank	Word	Freq.	%	Rank	Word	Freq.	%
981.	somnolence	48	0.01	982.	sterile	48	0.01
983.	stimulation	48	0.01	984.	vessels	48	0.01
985.	zestril	48	0.01	986.	arterial	47	0.01
987.	capsule	47	0.01	988.	carcinogenicity	47	0.01
989.	celebrex	47	0.01	990.	contact	47	0.01
991.	elevations	47	0.01	992.	excipients	47	0.01
993.	fall	47	0.01	994.	fatigue	47	0.01
995.	female	47	0.01	996.	hycamtin	47	0.01
997.	insomnia	47	0.01	998.	pharmacological	47	0.01
999.	salt	47	0.01	1000.	toremifene	47	0.01
1001.	value	47	0.01	1002.	allergy	46	0.01
1003.	assay	46	0.01	1004.	basis	46	0.01
1005.	consideration	46	0.01	1006.	criteria	46	0.01
1007.	differences	46	0.01	1008.	effectiveness	46	0.01
1009.	hypertensive	46	0.01	1010.	infant	46	0.01
1011.	initiation	46	0.01	1012.	instances	46	0.01
1013.	loratadine	46	0.01	1014.	metabolised	46	0.01
1015.	minutes	46	0.01	1016.	procedure	46	0.01
1017.	prothrombin	46	0.01	1018.	reducing	46	0.01
1019.	reproduction	46	0.01	1020.	saline	46	0.01
1021.	sinusitis	46	0.01	1022.	spiriva	46	0.01
1023.	suggest	46	0.01	1024.	tumours	46	0.01
1025.	anaphylactic	45	0.01	1026.	cmax	45	0.01
1027.	completely	45	0.01	1028.	desired	45	0.01
1029.	esmeron	45	0.01	1030.	fu	45	0.01
1031.	lactamase	45	0.01	1032.	megestrol	45	0.01
1033.	membrane	45	0.01	1034.	minor	45	0.01
1035.	monotherapy	45	0.01	1036.	occurrence	45	0.01
1037.	premature	45	0.01	1038.	pseudoephedrine	45	0.01
1039.	reduces	45	0.01	1040.	schedule	45	0.01
1041.	sickness	45	0.01	1042.	singulair	45	0.01
1043.	stored	45	0.01	1044.	streptococci	45	0.01
1045.	support	45	0.01	1046.	tamoxifen	45	0.01
1047.	wk	45	0.01	1048.	altered	44	0.01
1049.	applied	44	0.01	1050.	become	44	0.01
1051.	follow	44	0.01	1052.	hematologic	44	0.01
1053.	highly	44	0.01	1054.	hyperthyroidism	44	0.01
1055.	increasing	44	0.01	1056.	inhibace	44	0.01
1057.	microorganisms	44	0.01	1058.	phenytoin	44	0.01
1059.	piroxicam	44	0.01	1060.	receptors	44	0.01
1061.	rise	44	0.01	1062.	substance	44	0.01
1063.	tritace	44	0.01	1064.	unit	44	0.01
1065.	unlikely	44	0.01	1066.	weakness	44	0.01
1067.	acne	43	0.01	1068.	association	43	0.01
1069.	bile	43	0.01	1070.	bottle	43	0.01
1071.	capsules	43	0.01	1072.	dactinomycin	43	0.01
1073.	delayed	43	0.01	1074.	diagnosis	43	0.01
1075.	durogesic	43	0.01	1076.	factors	43	0.01
1077.	hemodialysis	43	0.01	1078.	intake	43	0.01
1079.	kaposi's	43	0.01	1080.	majority	43	0.01
1081.	mouthpiece	43	0.01	1082.	open	43	0.01
1083.	osteoarthritis	43	0.01	1084.	palpitations	43	0.01
1085.	pimecrolimus	43	0.01	1086.	potentiate	43	0.01
1087.	ratio	43	0.01	1088.	respirator	43	0.01
1089.	slowly	43	0.01	1090.	surgical	43	0.01

Rank	Word	Freq.	%	Rank	Word	Freq.	%
1091.	ug	43	0.01	1092.	would	43	0.01
1093.	airway	42	0.01	1094.	alkaline	42	0.01
1095.	anaemia	42	0.01	1096.	appears	42	0.01
1097.	artery	42	0.01	1098.	attacks	42	0.01
1099.	bp	42	0.01	1100.	cirrhosis	42	0.01
1101.	comparable	42	0.01	1102.	determine	42	0.01
1103.	diabetic	42	0.01	1104.	enhance	42	0.01
1105.	et	42	0.01	1106.	etc	42	0.01
1107.	fatal	42	0.01	1108.	interval	42	0.01
1109.	jaundice	42	0.01	1110.	joint	42	0.01
1111.	man	42	0.01	1112.	meropenem	42	0.01
1113.	minimum	42	0.01	1114.	mitomycin	42	0.01
1115.	mucosa	42	0.01	1116.	mutagenicity	42	0.01
1117.	plus	42	0.01	1118.	procef	42	0.01
1119.	pronounced	42	0.01	1120.	proteins	42	0.01
1121.	ranged	42	0.01	1122.	receive	42	0.01
1123.	regularly	42	0.01	1124.	remedies	42	0.01
1125.	statistically	42	0.01	1126.	tabs	42	0.01
1127.	thereafter	42	0.01	1128.	unknown	42	0.01
1129.	uterine	42	0.01	1130.	biliary	41	0.01
1131.	charcoal	41	0.01	1132.	container	41	0.01
1133.	double	41	0.01	1134.	eliminated	41	0.01
1135.	enhanced	41	0.01	1136.	flushing	41	0.01
1137.	handling	41	0.01	1138.	involving	41	0.01
1139.	ketoconazole	41	0.01	1140.	lithium	41	0.01
1141.	maternal	41	0.01	1142.	medicines	41	0.01
1143.	method	41	0.01	1144.	nasonex	41	0.01
1145.	pass	41	0.01	1146.	possibly	41	0.01
1147.	pyogenes	41	0.01	1148.	smooth	41	0.01
1149.	stada	41	0.01	1150.	steroid	41	0.01
1151.	tolerance	41	0.01	1152.	agonist	40	0.01
1153.	agonists	40	0.01	1154.	ampoules	40	0.01
1155.	aromasin	40	0.01	1156.	catarrhalis	40	0.01
1157.	causal	40	0.01	1158.	causes	40	0.01
1159.	cautions	40	0.01	1160.	cefuroxime	40	0.01
1161.	commonly	40	0.01	1162.	cramps	40	0.01
1163.	cross	40	0.01	1164.	death	40	0.01
1165.	fentanyl	40	0.01	1166.	forms	40	0.01
1167.	lowering	40	0.01	1168.	malignant	40	0.01
1169.	men	40	0.01	1170.	methods	40	0.01
1171.	morning	40	0.01	1172.	parent	40	0.01
1173.	secondary	40	0.01	1174.	stomach	40	0.01
1175.	swelling	40	0.01	1176.	tested	40	0.01
1177.	tremor	40	0.01	1178.	vertigo	40	0.01
1179.	yellow	40	0.01	1180.	added	39	0.01
1181.	aerosol	39	0.01	1182.	analysis	39	0.01
1183.	attained	39	0.01	1184.	compounds	39	0.01
1185.	corresponding	39	0.01	1186.	cumulative	39	0.01
1187.	dyspepsia	39	0.01	1188.	injections	39	0.01
1189.	instituted	39	0.01	1190.	intra	39	0.01
1191.	minoxidil	39	0.01	1192.	mutagenic	39	0.01
1193.	neuromuscular	39	0.01	1194.	neutrophil	39	0.01
1195.	nifedipin	39	0.01	1196.	nostril	39	0.01
1197.	product	39	0.01	1198.	recovery	39	0.01
1199.	red	39	0.01	1200.	show	39	0.01

Rank	Word	Freq.	%	Rank	Word	Freq.	%
1201.	swallowed	39	0.01	1202.	uncommon	39	0.01
1203.	withdrawn	39	0.01	1204.	zone	39	0.01
1205.	acetonide	38	0.01	1206.	anaphylaxis	38	0.01
1207.	aqueous	38	0.01	1208.	asthenia	38	0.01
1209.	circulation	38	0.01	1210.	complications	38	0.01
1211.	confusion	38	0.01	1212.	congestion	38	0.01
1213.	deterioration	38	0.01	1214.	digestive	38	0.01
1215.	dosages	38	0.01	1216.	drops	38	0.01
1217.	imatinib	38	0.01	1218.	letrozole	38	0.01
1219.	little	38	0.01	1220.	maois	38	0.01
1221.	megace	38	0.01	1222.	myelosuppression	38	0.01
1223.	necrosis	38	0.01	1224.	neuropathy	38	0.01
1225.	prostate	38	0.01	1226.	recovered	38	0.01
1227.	requiring	38	0.01	1228.	respond	38	0.01
1229.	rhythm	38	0.01	1230.	structure	38	0.01
1231.	sufficient	38	0.01	1232.	syringe	38	0.01
1233.	teratogenicity	38	0.01	1234.	wkly	38	0.01
1235.	anaphylactoid	37	0.01	1236.	angle	37	0.01
1237.	arava	37	0.01	1238.	atrial	37	0.01
1239.	bacteria	37	0.01	1240.	beta	37	0.01
1241.	ciclosporin	37	0.01	1242.	clearly	37	0.01
1243.	copd	37	0.01	1244.	enantone	37	0.01
1245.	hemolytic	37	0.01	1246.	ii	37	0.01
1247.	leukemia	37	0.01	1248.	medications	37	0.01
1249.	metoprolol	37	0.01	1250.	nasacort	37	0.01
1251.	nonsteroidal	37	0.01	1252.	respules	37	0.01
1253.	second	37	0.01	1254.	size	37	0.01
1255.	spontaneous	37	0.01	1256.	terminal	37	0.01
1257.	three	37	0.01	1258.	tinnitus	37	0.01
1259.	analgesic	36	0.01	1260.	anthracycline	36	0.01
1261.	anticoagulant	36	0.01	1262.	antineoplastics	36	0.01
1263.	aq	36	0.01	1264.	aqua	36	0.01
1265.	areas	36	0.01	1266.	cellulose	36	0.01
1267.	clear	36	0.01	1268.	conduction	36	0.01
1269.	consistent	36	0.01	1270.	delivery	36	0.01
1271.	diethelm	36	0.01	1272.	exposed	36	0.01
1273.	extensive	36	0.01	1274.	gain	36	0.01
1275.	induce	36	0.01	1276.	induction	36	0.01
1277.	lowest	36	0.01	1278.	monitor	36	0.01
1279.	nor	36	0.01	1280.	output	36	0.01
1281.	peptic	36	0.01	1282.	periods	36	0.01
1283.	probenecid	36	0.01	1284.	prostatic	36	0.01
1285.	provided	36	0.01	1286.	requires	36	0.01
1287.	revealed	36	0.01	1288.	theophylline	36	0.01
1289.	trimebutine	36	0.01	1290.	ulcerative	36	0.01
1291.	weekly	36	0.01	1292.	adjustments	35	0.01
1293.	adjuvant	35	0.01	1294.	aureus	35	0.01
1295.	beclomethasone	35	0.01	1296.	benefits	35	0.01
1297.	cml	35	0.01	1298.	coli	35	0.01
1299.	ear	35	0.01	1300.	end	35	0.01
1301.	faeces	35	0.01	1302.	influence	35	0.01
1303.	insufficient	35	0.01	1304.	lanoxin	35	0.01
1305.	malaise	35	0.01	1306.	monopril	35	0.01
1307.	mother	35	0.01	1308.	new	35	0.01
1309.	pharyngitis	35	0.01	1310.	place	35	0.01

Rank	Word	Freq.	%	Rank	Word	Freq.	%
1311.	products	35	0.01	1312.	ramipril	35	0.01
1313.	species	35	0.01	1314.	together	35	0.01
1315.	triamcinolone	35	0.01	1316.	usp	35	0.01
1317.	whose	35	0.01				

APPENDIX D

The General Service List (GSL) Found in the Pharmaceutical Corpus (lemmatized)

This frequency list contains 473 words found in the General Service List (GSL).

No.	Lemmas	Freq.	%	Words
1.	be	13,773	7.46	being(87),are(1503),am(3),is(4259),was(1003),were(826),been(1581)
2.	the	12,699	6.88	
3.	of	12,530	6.79	
4.	in	9,167	4.97	
5.	and	8,572	4.65	
6.	to	6,016	3.26	
7.	with	5,053	2.74	
8.	a	4,948	2.68	an(893)
9.	or	4,116	2.23	
10.	patient	3,790	2.05	patient's(118),patients(3250),patients(2)
11.	should	2,616	1.42	
12.	have	2,593	1.41	has(1030),having(23),had(169)
13.	for	2,524	1.37	
14.	may	2,059	1.12	
15.	use	1,902	1.03	used(654),using(187)
16.	not	1,860	1.01	
17.	treatment	1,677	0.91	treatments(21)
18.	as	1,642	0.89	
19.	by	1,535	0.83	
20.	effect	1,454	0.79	effected(1),effects(831)
21.	at	1,283	0.70	
22.	on	1,209	0.66	
23.	day	1,108	0.60	days(453)
24.	increase	1,096	0.59	increased(572),increases(122),increasing(44)
25.	it	1,031	0.56	its(295)
26.	study	1,014	0.55	studying(2),studied(64),studies(666)
27.	other	996	0.54	others(30)
28.	after	983	0.53	
29.	that	950	0.51	
30.	no	942	0.51	
31.	daily	924	0.50	
32.	hour	923	0.50	hours(138),hr(245),hrs(513)
33.	report	844	0.46	reported(657),reporting(4),reports(134)
34.	from	817	0.44	
35.	blood	786	0.43	
36.	see	736	0.40	seeing(1),seen(120)
37.	during	733	0.40	
38.	when	717	0.39	
39.	following	683	0.37	
40.	time	677	0.37	times(364),timing(3)
41.	recommend	660	0.36	recommends(1),recommended(655)
42.	this	655	0.35	
43.	include	637	0.35	included(61),includes(6),including(477)
44.	severe	636	0.34	
45.	than	618	0.33	
46.	these	613	0.33	
47.	can	586	0.32	cannot(54),could(72)
48.	give	566	0.31	given(464),gives(9),giving(28)

No.	Lemmas	Freq.	%	Words
49.	child	560	0.30	childn(100),children(441)
50.	treat	557	0.30	treated(505),treating(25)
51.	level	530	0.29	leveling(1),levels(410)
52.	case	525	0.28	cases(406),casing(2)
53.	there	524	0.28	
54.	which	521	0.28	
55.	high	517	0.28	higher(197),highest(59)
56.	observe	515	0.28	observed(508)
57.	week	488	0.26	weeks(292),wk(45)
58.	take	465	0.25	taken(268),takes(14),taking(124)
59.	result	460	0.25	results(152),resulting(78),resulted(55)
60.	also	455	0.25	
61.	receive	453	0.25	received(140),receiving(271)
62.	action	445	0.24	actions(230)
63.	year	438	0.24	year's(2),years(292),yr(89)
64.	low	430	0.23	lowered(6),lowering(40),lowers(10),lower(226),lowest(36)
65.	pack	426	0.23	packs(8),packing(396),packed(1)
66.	do	425	0.23	did(84),does(163),doing(3),done(18)
67.	test	423	0.23	testing(79),tests(142),tested(40)
68.	control	421	0.23	controlled(208),controlling(5),controls(20)
69.	solution	419	0.23	solutions(75)
70.	reduce	416	0.23	reduced(250),reduces(45),reducing(46)
71.	any	408	0.22	
72.	who	406	0.22	
73.	pressure	405	0.22	pressures(3)
74.	agent	387	0.21	agents(261)
75.	skin	387	0.21	
76.	contain	382	0.21	contained(7),containing(155),contains(189)
77.	minute	381	0.21	minutes(46),mins(1),min(306)
78.	cause	379	0.21	caused(138),causes(40),causing(27)
79.	only	359	0.19	
80.	decrease	353	0.19	decreased(159),decreases(52)
81.	month	351	0.19	months(276)
82.	risk	350	0.19	risks(21)
83.	show	349	0.19	showed(57),shown(230),shows(20),showing(3)
84.	change	345	0.19	changed(11),changes(172),changing(11),unchanged(84)
85.	rate	339	0.18	rated(2),rates(55)
86.	more	338	0.18	
87.	due	334	0.18	
88.	event	330	0.18	events(232)
89.	pain	327	0.18	pains(10)
90.	single	323	0.18	
91.	but	319	0.17	
92.	human	318	0.17	humans(81)
93.	must	317	0.17	
94.	once	308	0.17	
95.	caution	307	0.17	cautioned(3),cautions(40),cautious(4)
96.	body	306	0.17	bodies(1),body's(2)
97.	under	302	0.16	
98.	all	300	0.16	
99.	heart	296	0.16	
100.	presentation	291	0.16	presentations(6)
101.	associate	289	0.16	associated(288)
102.	usually	288	0.16	

No.	Lemmas	Freq.	%	Words
103.	within	284	0.15	
104.	trial	282	0.15	trials(222)
105.	discontinue	280	0.15	discontinued(186),discontinuing(25),discountinued(1)
106.	those	279	0.15	
107.	up	277	0.15	
108.	manufacturer	275	0.15	manufacturer's(1)
109.	such	274	0.15	
110.	activity	270	0.15	activities(27)
111.	however	267	0.14	
112.	life	267	0.14	
113.	therefore	267	0.14	
114.	some	266	0.14	
115.	woman	266	0.14	women(249)
116.	before	265	0.14	
117.	reduction	264	0.14	reductions(17)
118.	combination	262	0.14	combinations(5)
119.	each	260	0.14	
120.	failure	260	0.14	failures(1)
121.	relate	258	0.14	related(250),relating(4)
122.	weight	257	0.14	weights(7)
123.	great	255	0.14	greatest(5),greater(241)
124.	long	255	0.14	longer(59),longest(2)
125.	condition	254	0.14	conditions(177)
126.	most	250	0.14	
127.	know	247	0.13	knows(1),known(237)
128.	about	245	0.13	
129.	rarely	243	0.13	
130.	elderly	241	0.13	
131.	into	235	0.13	
132.	necessary	235	0.13	
133.	consider	233	0.13	considered(204),considering(6)
134.	equal	233	0.13	
135.	without	231	0.13	
136.	twice	229	0.12	
137.	well	228	0.12	
138.	active	226	0.12	
139.	follw	221	0.12	followed(117),follows(60)
140.	first	215	0.12	
141.	over	214	0.12	
142.	half	210	0.11	
143.	start	210	0.11	started(54),starting(81),starts(5)
144.	between	208	0.11	
145.	produce	205	0.11	produced(48),produces(27),producing(68)
146.	mean	204	0.11	means(22)
147.	since	204	0.11	
148.	group	202	0.11	groups(49)
149.	safety	200	0.11	
150.	animal	198	0.11	animals(85)
151.	possible	198	0.11	
152.	rare	198	0.11	
153.	side	198	0.11	sided(1),sides(7)
154.	table	198	0.11	
155.	class	195	0.11	classed(1),classes(4)
156.	both	191	0.10	

No.	Lemmas	Freq.	%	Words
157.	mild	191	0.10	
158.	because	190	0.10	
159.	small	186	0.10	smaller(65),smallest(4)
160.	every	184	0.10	
161.	store	184	0.10	stored(45),stores(3)
162.	sign	180	0.10	signs(163)
163.	subject	179	0.10	subjected(1),subjects(161)
164.	bleed	178	0.10	bleeding(168),bleeds(5),bleedings(1)
165.	content	178	0.10	contents(157)
166.	headache	177	0.10	headaches(11)
167.	water	177	0.10	
168.	value	176	0.10	values(129)
169.	one	173	0.09	
170.	until	168	0.09	
171.	avoid	167	0.09	avoids(1),avoided(67),avoiding(12)
172.	count	167	0.09	counting(1),counts(63)
173.	state	167	0.09	stated(6),states(20)
174.	effective	160	0.09	
175.	either	159	0.09	
176.	generally	156	0.08	
177.	although	155	0.08	
178.	history	155	0.08	
179.	base	154	0.08	based(134),bases(1)
180.	very	153	0.08	
181.	develop	152	0.08	developed(26),developing(27),develops(26)
182.	discontinuation	152	0.08	discontinuations(1)
183.	least	152	0.08	
184.	term	152	0.08	
185.	total	152	0.08	
186.	less	150	0.08	
187.	moderate	149	0.08	
188.	warning	149	0.08	warnings(88)
189.	continue	148	0.08	continued(93),continues(5),continuing(18)
190.	milk	147	0.08	
191.	repeat	143	0.08	repeated(119)
192.	drive	142	0.08	driven(5),driving(21)
193.	respectively	142	0.08	
194.	serious	141	0.08	
195.	compare	140	0.08	compared(124),comparing(14)
196.	they	139	0.08	
197.	common	138	0.07	
198.	fever	138	0.07	
199.	perform	138	0.07	performed(126),performing(1)
200.	while	138	0.07	
201.	limit	137	0.07	limits(11),limited(71),limiting(29)
202.	white	137	0.07	
203.	general	135	0.07	
204.	against	134	0.07	
205.	need	134	0.07	needed(58),needs(15)
206.	ability	133	0.07	abilities(3)
207.	dry	133	0.07	drying(2),dried(7)
208.	bone	132	0.07	bones(3)
209.	will	132	0.07	
210.	like	131	0.07	

No.	Lemmas	Freq.	%	Words
211.	operate	130	0.07	operated(2),operating(22)
212.	note	125	0.07	noted(63),notes(1)
213.	lung	124	0.07	lungs(22)
214.	per	124	0.07	
215.	course	123	0.07	courses(50)
216.	pen	123	0.07	
217.	machinery	122	0.07	
218.	price	122	0.07	
219.	immediately	121	0.07	
220.	usual	120	0.07	
221.	especially	117	0.06	
222.	post	116	0.06	
223.	block	115	0.06	blocking(61),blocked(1),blocks(4)
224.	coat	115	0.06	coated(102),coating(8),coatings(2)
225.	cough	115	0.06	coughing(12)
226.	loss	115	0.06	losses(6)
227.	measures	115	0.06	
228.	special	115	0.06	
229.	few	114	0.06	
230.	their	113	0.06	
231.	reach	112	0.06	reached(53),reaches(23),reaching(13)
232.	description	111	0.06	descriptions(2)
233.	determine	111	0.06	determined(60),determines(1),determining(8)
234.	local	111	0.06	
235.	resistant	111	0.06	
236.	according	109	0.06	
237.	bind	109	0.06	binding(86),binds(17)
238.	mouth	109	0.06	mouths(1)
239.	same	109	0.06	
240.	through	109	0.06	
241.	temperature	107	0.06	temperatures(13),temp(16)
242.	relief	106	0.06	
243.	alone	105	0.06	
244.	further	105	0.06	
245.	several	105	0.06	
246.	healthy	104	0.06	
247.	nursing	104	0.06	
248.	age	101	0.05	aged(4),ages(4),aging(1)
249.	where	101	0.05	
250.	below	100	0.05	
251.	mother	100	0.05	mother's(4),mothers(61)
252.	preparation	100	0.05	preparations(48)
253.	steady	99	0.05	
254.	even	98	0.05	evening(25)
255.	amount	97	0.05	amounting(1),amounts(33)
256.	expected	97	0.05	
257.	property	97	0.05	properties(94)
258.	line	96	0.05	
259.	provide	96	0.05	provided(36),providing(6),provides(28)
260.	flow	95	0.05	flows(2)
261.	make	95	0.05	makes(9),made(67),making(6)
262.	care	93	0.05	
263.	carefully	93	0.05	
264.	lead	93	0.05	leading(22),leads(16)

No.	Lemmas	Freq.	%	Words
265.	standard	92	0.05	
266.	upon	92	0.05	
267.	addition	91	0.05	additions(1)
268.	rapid	91	0.05	
269.	then	90	0.05	
270.	dog	89	0.05	dogs(72)
271.	improvement	89	0.05	improvements(3)
272.	depend	88	0.05	depends(12),depending(71)
273.	act	87	0.05	acting(48),acts(32)
274.	close	87	0.05	closed(34)
275.	become	86	0.05	becomes(31),becoming(9),became(2)
276.	light	86	0.05	
277.	dependent	85	0.05	
278.	out	85	0.05	
279.	suggest	85	0.05	suggested(12),suggesting(9),suggests(18)
280.	you	85	0.05	your(86)
281.	food	84	0.05	foods(1)
282.	type	84	0.05	types(17)
283.	meal	83	0.04	meals(52)
284.	excessive	82	0.04	
285.	man	82	0.04	men(40)
286.	management	82	0.04	
287.	careful	81	0.04	
288.	male	81	0.04	males(17)
289.	rapidly	81	0.04	
290.	thus	81	0.04	
291.	cold	80	0.04	
292.	extent	80	0.04	extents(1)
293.	present	80	0.04	
294.	prevent	80	0.04	prevented(7),preventing(9)
295.	protect	80	0.04	protected(17),protecting(2),protects(7)
296.	difference	79	0.04	differences(46)
297.	presence	79	0.04	
298.	unless	78	0.04	
299.	advance	77	0.04	advanced(75),advancing(1)
300.	likely	77	0.04	
301.	powder	77	0.04	
302.	two	77	0.04	
303.	attack	76	0.04	attacks(42)
304.	improve	76	0.04	improved(34),improving(7),improves(17)
305.	information	76	0.04	
306.	mainly	76	0.04	
307.	doctor	75	0.04	doctor's(7),doctors(1)
308.	medicine	75	0.04	medicines(41)
309.	product	75	0.04	product's(1),products(35)
310.	whether	75	0.04	
311.	cream	74	0.04	
312.	open	74	0.04	opened(18),opening(13)
313.	weekly	74	0.04	wkly(38)
314.	complete	73	0.04	completed(11)
315.	growth	73	0.04	
316.	substance	73	0.04	substances(29)
317.	development	72	0.04	
318.	direct	72	0.04	directed(10),directing(2)

No.	Lemmas	Freq.	%	Words
319.	number	72	0.04	numbers(3)
320.	population	72	0.04	populations(18)
321.	short	72	0.04	
322.	average	71	0.04	averaged(4),averages(8),averaging(1)
323.	upper	71	0.04	
324.	damage	70	0.04	damaged(3),damaging(1)
325.	resistance	70	0.04	resistance's (1)
326.	central	69	0.04	
327.	delay	69	0.04	delayed(43),delays(4)
328.	important	69	0.04	
329.	measure	69	0.04	measured(57)
330.	rabbit	69	0.04	rabbits(63)
331.	relationship	69	0.04	relationships(1)
332.	additional	68	0.04	
333.	early	68	0.04	
334.	progression	68	0.04	
335.	female	67	0.04	females(20)
336.	large	67	0.04	
337.	slow	67	0.04	slowed(3),slowing(2),slower(11)
338.	threatening	67	0.04	threatened(5)
339.	combine	66	0.04	combined(62),combines(1),combining(1)
340.	film	66	0.04	
341.	bottle	65	0.04	bottles(22)
342.	prevention	65	0.04	
343.	certain	64	0.03	
344.	closely	64	0.03	
345.	container	64	0.03	containers(23)
346.	many	64	0.03	
347.	described	63	0.03	
348.	possibility	63	0.03	
349.	room	63	0.03	
350.	advise	62	0.03	advised(59),advises(1)
351.	joint	62	0.03	joints(20)
352.	mark	62	0.03	marked(55),markings(1),marking(1),marks(1)
353.	salt	62	0.03	salts(15)
354.	slight	60	0.03	
355.	slightly	60	0.03	
356.	different	59	0.03	
357.	double	59	0.03	doubled(14),doubling(4)
358.	regular	59	0.03	
359.	fall	58	0.03	falling(3),falls(8)
360.	suitable	58	0.03	
361.	young	58	0.03	younger(30),youngest(1)
362.	above	57	0.03	
363.	clear	57	0.03	cleared(15),clearing(5),clears(1)
364.	inactive	57	0.03	
365.	eye	56	0.03	eyes(23)
366.	account	55	0.03	accounted(8),accounting(3),accounts(10)
367.	add	55	0.03	added(39),adding(4)
368.	fast	55	0.03	fasted(9),faster(4),fasting(28)
369.	market	55	0.03	marketed(2),marketing(52)
370.	order	55	0.03	
371.	support	55	0.03	supported(6),supporting(4)
372.	among	54	0.03	

No.	Lemmas	Freq.	%	Words
373.	brain	54	0.03	
374.	findings	54	0.03	
375.	shock	54	0.03	
376.	whole	54	0.03	
377.	reproductive	53	0.03	
378.	surface	53	0.03	surfaces(2)
379.	swallow	53	0.03	swallowed(39),swallowing(8)
380.	drop	52	0.03	drops(38),dropped(3)
381.	bound	51	0.03	
382.	continuous	51	0.03	
383.	essential	51	0.03	
384.	face	51	0.03	faces(1)
385.	sensitive	51	0.03	
386.	sometimes	51	0.03	
387.	supportive	51	0.03	
388.	cross	50	0.03	crosses(10)
389.	full	50	0.03	
390.	handle	50	0.03	handled(8),handling(41)
391.	taste	50	0.03	
392.	date	49	0.03	dates(1)
393.	free	49	0.03	
394.	often	49	0.03	
395.	various	49	0.03	
396.	absence	48	0.03	
397.	death	48	0.03	deaths(8)
398.	hair	48	0.03	
399.	influence	48	0.03	influenced(10),influences(3)
400.	so	48	0.03	
401.	consideration	47	0.03	considerations(1)
402.	measurement	47	0.03	measurements(15)
403.	stop	47	0.03	stopped(16),stopping(13),stops(2)
404.	basis	46	0.02	
405.	begin	46	0.02	began(1),beginning(31),begins(2),begun(6)
406.	desire	46	0.02	desired(45)
407.	effectiveness	46	0.02	
408.	exist	46	0.02	existing(30),exists(12)
409.	reproduction	46	0.02	
410.	completely	45	0.02	
411.	gain	45	0.02	gained(6),gaining(1),gains(2)
412.	sickness	45	0.02	
413.	highly	44	0.02	
414.	last	44	0.02	lasted(1),lasting(15)
415.	prepare	44	0.02	prepared(23),preparing(4)
416.	rise	44	0.02	
417.	unlikely	44	0.02	
418.	weakness	44	0.02	
419.	arm	43	0.02	arms(13)
420.	association	43	0.02	
421.	slowly	43	0.02	
422.	swell	43	0.02	swelling(40),swellings(2)
423.	would	43	0.02	
424.	comparable	42	0.02	
425.	complication	42	0.02	complications(38)
426.	pronounced	42	0.02	

No.	Lemmas	Freq.	%	Words
427.	regard	42	0.02	regarded(1),regarding(16),regards(4)
428.	regularly	42	0.02	
429.	smooth	42	0.02	smoothed(1)
430.	unknown	42	0.02	
431.	yellow	42	0.02	yellowing(2)
432.	keep	41	0.02	keeping(6),keeps(1)
433.	load	41	0.02	loaded(1),loading(32)
434.	pass	41	0.02	
435.	possibly	41	0.02	
436.	commonly	40	0.02	
437.	degree	40	0.02	degrees(9)
438.	morning	40	0.02	
439.	parent	40	0.02	
440.	recommendation	40	0.02	recommendations(31)
441.	secondary	40	0.02	
442.	stomach	40	0.02	
443.	angle	39	0.02	angles(2)
444.	employ	39	0.02	employed(31),employing(5)
445.	narrow	39	0.02	narrowing(12)
446.	observation	39	0.02	observations(5)
447.	red	39	0.02	
448.	size	39	0.02	sized(1),sizes(1)
449.	uncommon	39	0.02	
450.	check	38	0.02	checked(20),checking(3),checks(2)
451.	confusion	38	0.02	
452.	little	38	0.02	
453.	pump	38	0.02	pumping(1),pumps(4)
454.	shake	38	0.02	shaken(3),shaking(5)
455.	stage	38	0.02	stages(12),staging(3)
456.	carry	37	0.02	carrying(3),carried(27),carries(2)
457.	clearly	37	0.02	
458.	ear	37	0.02	ears(2)
459.	fail	37	0.02	failed(20),failing(4),fails(10)
460.	point	37	0.02	pointes(15),pointing(1),points(3)
461.	second	37	0.02	
462.	delivery	36	0.02	
463.	end	36	0.02	ending(1)
464.	extensive	36	0.02	
465.	model	36	0.02	models(24)
466.	monkey	36	0.02	monkeys(32)
467.	nor	36	0.02	
468.	except	35	0.02	excepting(1)
469.	list	35	0.02	listed(28),lists(3)
470.	new	35	0.02	
471.	place	35	0.02	
472.	together	35	0.02	
473.	whose	35	0.02	
Total		**184,514**	**100**	

APPENDIX E

The Academic Word List (AWL) Found in the Pharmaceutical Corpus (lemmatized)

This frequency list contains 113 lemmas.

No.	Lemmas	Freq.	%	Words
1.	administration	1.083	7.28	administrations(7)
2.	occur	864	5.81	occurred(156),occurring(55),occurs(140)
3.	concentration	622	4.18	concentrations(419)
4.	function	524	3.52	functions(23)
5.	adult	424	2.85	adults(256)
6.	require	396	2.66	required(198),requires(36),requiring(38)
7.	monitor	368	2.47	monitored(191),monitoring(141)
8.	indication	357	2.40	indications(337)
9.	indicate	306	2.06	indicated(159),indicates(48),indicating(28)
10.	approximately	305	2.05	
11.	range	267	1.80	ranged(42),ranges(14),ranging(23)
12.	potential	245	1.65	
13.	significant	243	1.63	
14.	data	231	1.55	
15.	incidence	203	1.37	incidences(16)
16.	inhibit	201	1.35	incidences(16)
17.	appropriate	197	1.32	
18.	maximum	189	1.27	
19.	elimination	185	1.24	
20.	establish	181	1.22	established(164),establishing(9)
21.	induce	179	1.20	induced(116),induces(16),inducing(11)
22.	inhibitor	176	1.18	inhibitors(102)
23.	phase	170	1.14	phased(1),phases(3)
24.	evidence	164	1.10	evidenced(4)
25.	duration	157	1.06	durations(2)
26.	prior	155	1.04	
27.	initial	149	1.00	
28.	similar	149	1.00	
29.	maintenance	134	0.90	
30.	physician	134	0.90	physicians(14),physician's(10)
31.	reversible	134	0.90	
32.	adjustment	133	0.89	adjustments(35)
33.	affect	131	0.88	affected(56),affecting(10),affects(13)
34.	annual	128	0.86	
35.	interval	128	0.86	intervals(86)
36.	initiate	122	0.82	initiated(88),initiates(1),initiating(19)
37.	insufficiency	120	0.81	
38.	volume	118	0.79	volumes(5)
39.	site	117	0.79	sites(26)
40.	specific	116	0.78	
41.	inhibition	115	0.77	
42.	grade	113	0.76	grades(6),grading(1)
43.	obtain	109	0.73	obtained(86),obtaining(2)
44.	area	108	0.73	areas(36)
45.	individual	106	0.71	individuals(32)
46.	medical	104	0.70	
47.	adequate	103	0.69	

No.	Lemmas	Freq.	%	Words
48.	adjust	103	0.69	adjusted(82),adjusting(3)
49.	significantly	101	0.68	
50.	enhance	98	0.66	enhanced(41),enhancing(2),enhances(13)
51.	exceed	98	0.66	exceeded(10),exceeding(22),exceeds(10)
52.	compound	96	0.65	compounds(39)
53.	subsequent	96	0.65	
54.	factor	91	0.61	factors(43)
55.	maintain	91	0.61	maintained(54),maintaining(7)
56.	evaluate	86	0.58	evaluated(65),evaluating(4)
57.	isolate	86	0.58	isolated(81),isolates(3)
58.	involve	82	0.55	involved(22),involing(1),involves(14),involving(41)
59.	method	81	0.54	methods(40)
60.	via	80	0.54	
61.	major	75	0.50	
62.	medicine	75	0.50	medicines(41)
63.	volunteer	75	0.50	volunteers(73)
64.	alter	74	0.50	altered(44),altering(3),alters(3)
65.	mechanism	74	0.50	mechanisms(18)
66.	vision	74	0.50	
67.	conducted	71	0.48	
68.	parameter	71	0.48	parameters(69)
69.	cycle	66	0.44	cycles(32)
70.	instruction	66	0.44	instructions(59)
71.	primary	66	0.44	
72.	concurrent	61	0.41	
73.	primarily	61	0.41	
74.	route	61	0.41	routes(11)
75.	schedule	61	0.41	scheduled(5),schedules(11)
76.	detect	58	0.39	detected(53)
77.	apparent	56	0.38	
78.	initially	56	0.38	
79.	remove	56	0.38	removed(25),removing(5)
80.	potentially	54	0.36	
81.	contact	51	0.34	contacts(4)
82.	correspond	51	0.34	corresponding(39),corresponds(10)
83.	recover	51	0.34	recovered(38),recovers(2),recovering(1)
84.	visual	51	0.34	
85.	instance	50	0.34	instances(46)
86.	accompany	48	0.32	accompanying(10),accompanied(34)
87.	eliminate	48	0.32	eliminated(41),eliminates(1),eliminating(2)
88.	occurrence	48	0.32	occurrences(3)
89.	technique	48	0.32	techniques(27)
90.	initiation	46	0.31	
91.	minor	45	0.30	
92.	institute	44	0.30	instituted(39),institutes(1)
93.	device	43	0.29	devices(12)
94.	exclude	43	0.29	excluded(28),excluding(7)
95.	majority	43	0.29	
96.	ratio	43	0.29	
97.	minimum	42	0.28	
98.	plus	42	0.28	
99.	reveal	42	0.28	revealed(36)
100.	attain	41	0.28	attained(39),attaining(1)
101.	exhibit	41	0.28	exhibited(12),exhibiting(4),exhibits(15)

No.	Lemmas	Freq.	%	Words
102.	recovery	40	0.27	recoveries(1)
103.	analysis	39	0.26	
104.	estimate	38	0.26	estimated(18),estimates(11)
105.	structure	38	0.26	
106.	sufficient	38	0.26	
107.	assess	37	0.25	assessing(8),assessed(18)
108.	exposed	37	0.25	exposing(1)
109.	consistent	36	0.24	
110.	induction	36	0.24	
111.	minimise	36	0.24	minimised(2),minimising(2),minimized(8),minimizes(2),minimize(18),minimizing(1)
112.	insufficient	35	0.24	
113.	label	35	0.24	labeled(5),labeling(2),labelled(5)
Total		**14,869**	**100**	

APPENDIX F

The Technical Vocabulary List Found in the Pharmaceutical Corpus (lemmatized)

This frequency list contains 404 lemmas.

No.	Lemmas	Freq.	%	Words
1.	dose	2,866	6.00	dosed(16),doses(830),dosing(126)
2.	drug	1,358	2.84	drug's(6),drugs(433)
3.	therapy	1,170	2.45	therapies(10)
4.	dosage	1,076	2.25	dosages(38),dosaging(1)
5.	tablet	986	2.07	tabs(42),tablets(234),tablet's(1),tab(472)
6.	adverse	790	1.65	
7.	renal	779	1.63	
8.	administer	720	1.51	administered(654),administering(34),administrered(2)
9.	clinical	689	1.44	
10.	plasma	620	1.30	
11.	oral	567	1.19	
12.	symptom	530	1.11	symptoms(505)
13.	pregnancy	501	1.05	pregnancies(5)
14.	infusion	464	0.97	infusions(34)
15.	precaution	462	0.97	precautions(449)
16.	acute	408	0.85	
17.	hepatic	395	0.83	
18.	serum	393	0.82	
19.	liver	369	0.77	
20.	disorder	326	0.68	disorders(256)
21.	corticosteroid	321	0.67	corticosteroids(188)
22.	impairment	318	0.67	
23.	contraindication	310	0.65	contraindications(53)
24.	hypersensitivity	294	0.62	
25.	cancer	286	0.60	cancers(17)
26.	cardiac	284	0.59	
27.	toxicity	274	0.57	
28.	impair	270	0.57	impaired(203),impairs(3)
29.	metabolite	258	0.54	metabolites(155)
30.	chronic	250	0.52	
31.	pharmacokinetics	247	0.52	
32.	overdosage	242	0.51	
33.	gastrointestinal	238	0.50	
34.	breast	236	0.49	
35.	capsule	230	0.48	capsules(43),caps(27),cap(113)
36.	hypertension	228	0.48	
37.	sodium	228	0.48	
38.	nausea	226	0.47	
39.	clearance	223	0.47	
40.	cefaclor	212	0.44	
41.	tumour	209	0.44	tumours(46),tumors(61),tumor(52)
42.	creatinine	199	0.42	
43.	lactation	199	0.42	
44.	asthma	198	0.41	
45.	therapeutic	197	0.41	
46.	concomitant	191	0.40	
47.	nasal	189	0.40	
48.	hypotension	188	0.39	

No.	Lemmas	Freq.	%	Words
49.	diarrhea	185	0.39	diarrhoea(71)
50.	allergic	184	0.39	
51.	acid	182	0.38	acids(6)
52.	tract	182	0.38	tracts(6)
53.	vial	180	0.38	vials(25)
54.	fluid	179	0.37	fluids(25)
55.	diuretic	178	0.37	diuretics(97)
56.	tissue	178	0.37	tissues(69)
57.	pregnant	176	0.37	
58.	edema	176	0.37	oedema(74),oedemas(6)
59.	disturbance	174	0.36	disturbances(148)
60.	syndrome	172	0.36	syndromes(7)
61.	rash	171	0.36	rashes(28)
62.	excreted	169	0.35	
63.	clinically	167	0.35	
64.	metabolism	167	0.35	
65.	dizziness	166	0.35	dizziness(165)
66.	urine	166	0.35	
67.	in vitro	165	0.35	
68.	pneumoniae	164	0.34	pneumonia(70)
69.	urinary	162	0.34	
70.	antibiotic	161	0.34	antibiotics(105)
71.	laboratory	154	0.32	laboratories(2),laboratory's(4)
72.	peripheral	154	0.32	
73.	discontinuation	152	0.32	discontinuations(1)
74.	inflammatory	149	0.31	
75.	syrup	148	0.31	syr(80)
76.	muscle	145	0.30	muscles(10)
77.	spray	145	0.30	sprayed(4),sprays(23)
78.	respiratory	144	0.30	
79.	enzyme	140	0.29	enzymes(49)
80.	salbutamol	140	0.29	
81.	myocardial	137	0.29	
82.	orally	137	0.29	
83.	protein	136	0.28	proteins(42)
84.	streptococcus	136	0.28	streptococci(45)
85.	potassium	135	0.28	
86.	physician	134	0.28	physicians(14),physician's(10)
87.	clarithromycin	133	0.28	
88.	storage	133	0.28	
89.	placebo	132	0.28	
90.	efficacy	130	0.27	
91.	tachycardia	129	0.27	
92.	inhalation	127	0.27	
93.	inhale	127	0.27	inhaled(124),inhales(1)
94.	medication	126	0.26	medications(37)
95.	multidose	125	0.26	
96.	excretion	121	0.25	
97.	infant	120	0.25	infants(74)
98.	insufficiency	120	0.25	
99.	intron	120	0.25	
100.	neutropenia	119	0.25	neutropenias(1)
101.	kidney	118	0.25	kidneys(28)
102.	anti	116	0.24	

No.	Lemmas	Freq.	%	Words
103.	digoxin	116	0.24	
104.	platelet	116	0.24	platelets(24)
105.	receptor	116	0.24	receptors(44)
106.	angina	114	0.24	
107.	transient	114	0.24	
108.	auc	112	0.23	aucs(3)
109.	thrombocytopenia	112	0.23	thrombocytopenias(1)
110.	anaemia	110	0.23	anemia(68)
111.	coronary	110	0.23	
112.	hepatitis	110	0.23	
113.	pharmacology	110	0.23	
114.	cisplatin	107	0.22	
115.	kaposi's sarcoma	107	0.22	kaposi's sarcomas(7)
116.	regimen	106	0.22	regimens(28)
117.	cardiovascular	104	0.22	
118.	antihypertensive	103	0.22	antihypertensives(12)
119.	carcinoma	102	0.21	carcinomas(6)
120.	anticoagulant	101	0.21	anticoagulants(65)
121.	susceptible	101	0.21	
122.	organism	100	0.21	organisms(77)
123.	gastric	95	0.20	
124.	toxic	95	0.20	
125.	glucose	94	0.20	
126.	rhinitis	94	0.20	
127.	symptomatic	94	0.20	
128.	ventricular	94	0.20	
129.	concomitantly	93	0.19	
130.	susceptibility	93	0.19	
131.	arthritis	92	0.19	
132.	flush	92	0.19	flushed(6),flushes(29),flushing(41)
133.	calcium	91	0.19	
134.	chemotherapy	91	0.19	chemotherapies(1)
135.	dilute	91	0.19	diluted(73),diluting(4)
136.	overdose	91	0.19	overdoses(10)
137.	pulmonary	91	0.19	
138.	marrow	90	0.19	
139.	ace	89	0.19	
140.	blocker	89	0.19	blockers(58)
141.	median	89	0.19	
142.	dysfunction	88	0.18	
143.	propionate	88	0.18	
144.	pruritus	88	0.18	
145.	colitis	87	0.18	
146.	fluticasone	87	0.18	
147.	urticaria	87	0.18	
148.	hydrochloride	86	0.18	
149.	abdominal	85	0.18	
150.	gabapentin	85	0.18	
151.	chloride	83	0.17	chlorides(2)
152.	diclofenac	83	0.17	
153.	dilution	83	0.17	dilutions(3)
154.	dobutamine	82	0.17	
155.	pharmacokinetic	82	0.17	
156.	retard	81	0.17	retarded(3)

No.	Lemmas	Freq.	%	Words
157.	seizure	81	0.17	seizures(69)
158.	surgery	81	0.17	surgeries(2)
159.	bioavailability	80	0.17	
160.	cephalosporin	80	0.17	cephalosporins(49)
161.	pediatric	80	0.17	
162.	adrenal	79	0.17	
163.	potentiate	79	0.17	potentiated(22),potentiates(11),potentiating(3)
164.	arrhythmia	78	0.16	arrhythmias(59)
165.	contraindicated	78	0.16	
166.	glaucoma	78	0.16	
167.	parenteral	78	0.16	
168.	dyspnea	77	0.16	dyspnoea(19)
169.	membrane	77	0.16	membranes(32)
170.	vascular	77	0.16	
171.	dermatitis	76	0.16	
172.	inhaler	76	0.16	inhalers(3)
173.	influenzae	75	0.16	
174.	ingredient	75	0.16	ingredients(50)
175.	prescribe	75	0.16	prescribed(55),prescribes(2),prescribing(13)
176.	baseline	74	0.16	
177.	bronchospasm	74	0.16	bronchospasms(1)
178.	haemophilus	74	0.16	
179.	interferon	74	0.16	interferon's(1)
180.	acetate	73	0.15	
181.	diabetes	73	0.15	
182.	nervous	73	0.15	
183.	paclitaxel	73	0.15	
184.	leukaemia	72	0.15	leukaemias(4),leukemias(3),leukemia(37)
185.	salmeterol	72	0.15	
186.	warfarin	72	0.15	
187.	alcohol	71	0.15	
188.	ampoule	71	0.15	ampoules(40)
189.	intestinal	71	0.15	
190.	doxorubicin	70	0.15	
191.	onset	70	0.15	
192.	budesonide	69	0.14	
193.	ovarian	69	0.14	
194.	quinolone	69	0.14	quinolones(57)
195.	teratogenic	69	0.14	
196.	antineoplastic	67	0.14	antineoplastics(36)
197.	bronchial	67	0.14	
198.	metastatic	67	0.14	
199.	palpitation	67	0.14	palpitations(43)
200.	assay	66	0.14	assayed(2),assaying(2),assays(16)
201.	cerebral	66	0.14	
202.	dextrose	66	0.14	
203.	sulfate	66	0.14	sulfated(1)
204.	dialysis	65	0.14	
205.	pectoris	65	0.14	
206.	secretion	65	0.14	secretions(8)
207.	fertility	64	0.13	
208.	potent	64	0.13	
209.	bradycardia	63	0.13	
210.	constipation	63	0.13	

No.	Lemmas	Freq.	%	Words
211.	maleate	63	0.13	
212.	preclinical	63	0.13	
213.	antacid	62	0.13	antacids(52)
214.	erythema	62	0.13	
215.	ulceration	62	0.13	ulcerations(2)
216.	disk	61	0.13	disks(5)
217.	fetus	61	0.13	fetuses(7)
218.	irritation	61	0.13	irritations(1)
219.	lesion	61	0.13	lesions(54)
220.	leukopenia	61	0.13	
221.	metabolic	60	0.13	
222.	rheumatoid	60	0.13	
223.	anorexia	59	0.12	
224.	doxazosin	59	0.12	
225.	drowsiness	59	0.12	
226.	intravenous	59	0.12	
227.	vivo	59	0.12	
228.	alopecia	58	0.12	
229.	angioedema	58	0.12	
230.	intranasal	58	0.12	
231.	magnesium	58	0.12	
232.	toxicology	58	0.12	
233.	bronchitis	57	0.12	
234.	heparin	56	0.12	
235.	isotretinoin	56	0.12	
236.	oxygen	56	0.12	
237.	prophylaxis	56	0.12	
238.	soluble	56	0.12	
239.	ulcer	56	0.12	
240.	usage	56	0.12	
241.	exacerbation	55	0.12	exacerbations(31)
242.	acne	54	0.11	acnes(11)
243.	cimetidine	54	0.11	
244.	excipient	54	0.11	excipients(47)
245.	adrenergic	53	0.11	adrenergics(1)
246.	diltiazem	53	0.11	
247.	docetaxel	53	0.11	
248.	penicillin	53	0.11	
249.	platinum	53	0.11	
250.	staphylococcus	53	0.11	
251.	synthesis	53	0.11	
252.	blurred	52	0.11	blurring(2)
253.	congestive	52	0.11	
254.	cytochrome	52	0.11	
255.	diameter	52	0.11	diameters(20)
256.	inflammation	52	0.11	inflammations(2)
257.	lamivudine	52	0.11	
258.	metabolise	52	0.11	metabolised(46),metabolising(1),metabolizing(3)
259.	supervision	52	0.11	
260.	angiotensin	51	0.11	
261.	neutrophil	51	0.11	neutrophils(12)
262.	antagonist	50	0.10	antagonists(26)
263.	anticholinergic	50	0.10	anticholinergics(20)
264.	fetal	50	0.10	

No.	Lemmas	Freq.	%	Words
265.	lactamase	50	0.10	lactamases(5)
266.	methotrexate	50	0.10	
267.	bilirubin	49	0.10	
268.	circulatory	49	0.10	
269.	diabetic	49	0.10	diabetics(7)
270.	hypertensive	49	0.10	hypertensives(3)
271.	anthracycline	48	0.10	anthracyclines(12)
272.	bacterial	48	0.10	
273.	expectorant	48	0.10	
274.	prostaglandin	48	0.10	prostaglandins(15)
275.	somnolence	48	0.10	
276.	sterile	48	0.10	
277.	stimulation	48	0.10	
278.	allergy	47	0.10	allergies(1)
279.	arterial	47	0.10	
280.	carcinogenicity	47	0.10	
281.	fatigue	47	0.10	
282.	insomnia	47	0.10	
283.	pharmacological	47	0.10	
284.	loratadine	46	0.10	
285.	microorganism	46	0.10	microorganisms(44)
286.	prothrombin	46	0.10	
287.	saline	46	0.10	
288.	sinusitis	46	0.10	
289.	tremor	46	0.10	tremors(6)
290.	anaphylactic	45	0.09	
291.	megestrol	45	0.09	
292.	monotherapy	45	0.09	
293.	neuropathy	45	0.09	neuropathies(7)
294.	premature	45	0.09	
295.	pseudoephedrine	45	0.09	
296.	tamoxifen	45	0.09	
297.	analgesic	44	0.09	analgesics(8)
298.	cramp	44	0.09	cramping(2),cramps(40)
299.	discard	44	0.09	discarded(34)
300.	hematologic	44	0.09	
301.	hyperthyroidism	44	0.09	
302.	opioid	44	0.09	opioids(15)
303.	phenytoin	44	0.09	
304.	piroxicam	44	0.09	
305.	bile	43	0.09	
306.	dactinomycin	43	0.09	
307.	diagnosis	43	0.09	
308.	hemodialysis	43	0.09	
309.	intake	43	0.09	
310.	mouthpiece	43	0.09	
311.	osteoarthritis	43	0.09	
312.	respirator	43	0.09	
313.	surgical	43	0.09	
314.	alkaline	42	0.09	
315.	artery	42	0.09	
316.	bronchodilator	42	0.09	bronchodilators(13)
317.	cirrhosis	42	0.09	
318.	fatal	42	0.09	

No.	Lemmas	Freq.	%	Words
319.	jaundice	42	0.09	
320.	mitomycin	42	0.09	
321.	mucosa	42	0.09	
322.	mutagenicity	42	0.09	
323.	nostril	42	0.09	nostrils(3)
324.	respule	42	0.09	respules(37)
325.	uterine	42	0.09	
326.	biliary	41	0.09	
327.	charcoal	41	0.09	
328.	glucocorticoid	41	0.09	glucocorticoids(19)
329.	ketoconazole	41	0.09	
330.	lithium	41	0.09	
331.	maternal	41	0.09	
332.	paresthesia	41	0.09	paresthesias(8)
333.	pyogenes	41	0.09	
334.	steroid	41	0.09	
335.	tolerance	41	0.09	
336.	catarrhalis	40	0.08	
337.	causal	40	0.08	
338.	cefuroxime	40	0.08	
339.	fentanyl	40	0.08	
340.	malignant	40	0.08	
341.	manifestation	40	0.08	manifestations(34)
342.	trimester	40	0.08	trimesters(7)
343.	vertigo	40	0.08	
344.	aerosol	39	0.08	
345.	breastfeed	39	0.08	breastfeeding(32)
346.	cumulative	39	0.08	
347.	dyspepsia	39	0.08	
348.	haemorrhage	39	0.08	haemorrhages(5)
349.	minoxidil	39	0.08	
350.	mutagenic	39	0.08	
351.	neuromuscular	39	0.08	
352.	syringe	39	0.08	syringes(1)
353.	zone	39	0.08	
354.	acetonide	38	0.08	
355.	anaphylaxis	38	0.08	
356.	antidote	38	0.08	antidotes(11)
357.	antihistamine	38	0.08	antihistamines(30)
358.	aqueous	38	0.08	
359.	asthenia	38	0.08	
360.	circulation	38	0.08	
361.	congestion	38	0.08	
362.	convulsion	38	0.08	convulsions(34)
363.	depressant	38	0.08	depressants(31)
364.	deterioration	38	0.08	
365.	digestive	38	0.08	
366.	myelosuppression	38	0.08	
367.	necrosis	38	0.08	
368.	nitrate	38	0.08	nitrates(17)
369.	precipitate	38	0.08	precipitated(8),precipitates(3),precipitating(3)
370.	prostate	38	0.08	
371.	rhythm	38	0.08	
372.	sympathomimetic	38	0.08	sympathomimetics(18)

No.	Lemmas	Freq.	%	Words
373.	teratogenicity	38	0.08	
374.	activated	37	0.08	activates(3),activating(5)
375.	anaphylactoid	37	0.08	
376.	antimicrobial	37	0.08	antimicrobials(4)
377.	atrial	37	0.08	
378.	bacteria	37	0.08	
379.	hemolytic	37	0.08	
380.	metoprolol	37	0.08	
381.	nonsteroidal	37	0.08	
382.	spontaneous	37	0.08	
383.	tinnitus	37	0.08	
384.	transaminase	37	0.08	transaminases(26)
385.	aqua	36	0.08	
386.	cellulose	36	0.08	
387.	peptic	36	0.08	
388.	perforation	36	0.08	perforations(2)
389.	probenecid	36	0.08	
390.	prostatic	36	0.08	
391.	sinus	36	0.08	sinuses(5)
392.	theophylline	36	0.08	
393.	ulcerative	36	0.08	
394.	adjuvant	35	0.07	
395.	antibacterial	35	0.07	antibacterials(2)
396.	aureus	35	0.07	
397.	beclomethasone	35	0.07	
398.	coli	35	0.07	
399.	faeces	35	0.07	
400.	malaise	35	0.07	
401.	pharyngitis	35	0.07	
402.	ramipril	35	0.07	
403.	species	35	0.07	
404.	triamcinolone	35	0.07	
Total		**48,159**	**100.00**	

APPENDIX G

List of Abbreviations Found in the Pharmaceutical Corpus

No.	Abbreviations	Freq.	%	Full Words
1.	mg	2,685	22.83	Milligram,
2.	ml	936	7.96	Milliliter
3.	kg	549	4.67	Kilogram
4.	eg	513	4.36	For example (Latin : exempli gratia)
5.	mcg	484	4.12	Microgram
6.	tab	472	4.01	Tablet
7.	iv	453	3.85	Intravenous
8.	c	309	2.63	Celsius/centigrade
9.	min	306	2.60	Minute, minimum
10.	hr(s)	245	2.08	Hour/hours
11.	l	216	1.84	Liter
12.	g	189	1.61	Gram
13.	w/	170	1.45	With
14.	p	164	1.39	Pressure
15.	s	143	1.22	Staphylococcus
16.	mims	139	1.18	Monthly Index of Medical Specialties
17.	mr	136	1.16	Milliroentgen
18.	hcl	120	1.02	Hydrochloride
19.	cap(s)	113	0.96	Capsule(s)
20.	i	108	0.92	1
21.	nsaids	107	0.91	Non-Steroidal Anti-Inflammatory Drug(s)
22.	tsp	106	0.90	Teaspoon / teaspoonful
23.	gi	103	0.88	Gastrointestinal
24.	childn	100	0.85	Children
25.	im	97	0.82	Intramuscular
26.	ace	89	0.76	Angiotensin Converting Enzyme
27.	yr	89	0.76	Year
28.	cns	88	0.75	Central nervous system
29.	h	86	0.73	Hour(Latin : hora)
30.	sc	86	0.73	Subcutaneous/scandium
31.	ng	84	0.71	Nanogram
32.	m	83	0.71	Thousand (Latin: mille), meter, mole
33.	syr	80	0.68	Syrup
34.	amp	76	0.65	Ampoules
35.	mm	76	0.65	Millimeter
36.	cr	69	0.59	Creatinine
37.	maoi(s)	68	0.58	Monoamine Oxidase Inhibitor(s)
38.	ph	66	0.56	Potential of Hydrogen
39.	tid	64	0.54	Three time a day (Latin : ter in die)
40.	miu	59	0.50	Million international units
41.	sp	57	0.48	Spoon
42.	mic	56	0.48	Minimum Inhibitory Concentration
43.	aids	55	0.47	Acquired Immuno-Deficiency Syndrome
44.	ie	54	0.46	That is (Latin : id est)
45.	cox	53	0.45	Cyclooxygenase
46.	dna	51	0.43	Deoxyribonucleic acid
47.	v	51	0.43	Against, in contrast to (Latin : versus)
48.	iii	50	0.43	3
49.	al	49	0.42	Artificial Insemination

No.	Abbreviations	Freq.	%	Full Words
50.	f	49	0.42	Fahrenheit
51.	cmax	45	0.38	Maximum Plasma Concentration
52.	wk	45	0.38	Week
53.	µg	43	0.37	Microgram
54.	bp	42	0.36	Base pair
55.	et al	42	0.36	And the others (Latin : et alii)
56.	etc	42	0.36	And other things(Latin: Et ectera)
57.	tabs	42	0.36	Tablets
58.	wkly	38	0.32	weekly
59.	ii	37	0.31	2
60.	aq	36	0.31	Aqueous
61.	cml	35	0.30	Chronic myelocytic leukemia/Chronic myelogenous leukemia/Chronic
62.	vs	33	0.28	Against, in contrast to (Latin : versus)
63.	bid	33	0.28	Twice A Day(Latin: bis in die)
64.	ecg	32	0.27	Electrocardiogram
65.	actt	32	0.27	American Type Culture Collection
66.	inj	31	0.26	Injection
67.	qid	30	0.26	Four time a day (Latin : quater in die)
68.	resp	30	0.26	Respiratory
69.	wt	29	0.25	Weight
70.	csf	25	0.21	Cerebro spinal fluid
71.	dl	24	0.20	Deciliter
72.	mth	24	0.20	Mouth
73.	hiv	23	0.20	Human immumodeficiency virus
74.	hrly	23	0.20	Hourly
75.	ks	23	0.20	Kaposi's sarcoma
76.	ast	21	0.18	Aspartate amino transferase
77.	hg	20	0.17	Mercury
78.	sgpt	19	0.16	Serum Glutamic Pyruvic Transaminase
79.	susp	19	0.16	Suspension
80.	ca	18	0.15	Calcium
81.	il	18	0.15	Interleukin
82.	hbr	17	0.14	Hydrobromide
83.	hpa	17	0.14	Hypothalamic-pituitary-adrenal
84.	gal	16	0.14	Galactose, gallon
85.	temp	16	0.14	Temperature
86.	tmax	16	0.14	Time to Maximum Plasma Concentration
87.	spp	15	0.13	Species
88.	nsclc	13	0.11	Non Small Cell Lung Cancer
89.	tsh	8	0.07	Thyroid-Stimulating Hormone
90.	azt	7	0.06	Azido-Thymidine (AIDS drug)
91.	clcr	7	0.06	Creatinine Clearance
92.	dpd	6	0.05	Dihydropyrimidine Dehydrogenase
93.	lh	6	0.05	Luteinizing Hormone
94.	tbsp	6	0.05	Tablespoonful
95.	exp	5	0.04	Expiration date
Total		**11,760**	**100.00**	

APPENDIX H

Number of Patients according to 21 Groups from Heath Service Units, Ministry of Public Health

Diseases	Number of Patients
Diseases of the respiratory system	25,199,868
Diseases of the digestive system	13,705,523
Diseases of the musculoskeletal system and connective tissue	8,563,866
Diseases of the skin and subcutaneous tissue	6,089,359
Diseases of the circulatory system	5,925,888

APPENDIX I

Number of Death according to the ICD Mortality Tabulation List 1, the 10th Revision

Diseases	Number of Death
Malignant neoplsm and all forms	39,480
Disease of circulatory system	32,331
Pneumonia and other disease of Lung	8,334
Nephritis, Nephrotic syndrome and nephrosis	8,156
Disease of liver and pancreases	6,110

BIOGRAPHY

Name: Miss Phatrinthorn Pongnan

Date of Birth: 16 December 1979

Place of Birth: Buriram, Thailand

Education:

2001	Bachelor of Education (Thai/English), Srinakharinwirot University
2004	Bachelor of Arts (English/Mass Communication), Ramkhamhaeng University
2006	Master of Arts (Applied Linguistics-English for Science and Technology), King MongKut's Institute of Technology Ladkrabang

Work Experience:

2001- 2003	Teaching Thai as a foreign language at SMITH International School of Languages.
2003-2005	Secretary to JICA EXPERT (Japan International Cooperation Agency) Computer Administration Center, Department of Employment, Ministry of Labour.
2006- present	Labour Technical Officer, Office of Inspector, Department of Employment, Ministry of Labour.