THESIS APPROVAL

GRADUATE SCHOOL, KASETSART UNIVERSITY

Master of Arts (English for Specific Purposes)

DEGREE

English for Specific Purposes

FIELD

Foreign Languages

DEPARTMENT

TITLE: Collocations of Keywords Found in Insurance Research Articles: A Corpus-Based Analysis

NAME: Mr. Tanu Khamphairoh

THIS THESIS HAS BEEN ACCEPTED BY

THESIS ADVISOR

(Mr. Montri Tangpijaikul, Ph.D.)

DEPARTMENT HEAD

(Assistant Professor Nongnuch Sriussadaporn, Ph.D.)

APPROVED BY THE GRADUATE SCHOOL ON

DEAN

(Associate Professor Gunjana Theeragool, D.Agr.)

THESIS

COLLOCATIONS OF KEYWORDS FOUND IN INSURANCE RESEARCH ARTICLES: A CORPUS-BASED ANALYSIS

TANU KHAMPHAIROH

A Thesis Submitted in Partial Fulfillment of

the Requirements of the Degree of

Master of Arts (English for Specific Purposes)

Graduate School, Kasetsart University

2012

Tanu Khamphairoh 2012: Collocations of Keywords Found in Insurance Research Articles: A Corpus-Based Analysis. Master of Arts (English for Specific Purposes), Major Field: English for Specific Purposes, Department of Foreign Languages. Thesis Advisor: Mr. Montri Tangpijaikul, Ph.D. 105 pages.

The objectives of this study are to identify keywords most frequently used in insurance research articles, and to explore the patterns of collocations of the keywords used in insurance research articles.

One hundred and fifty-five research articles from the Journal of Risk and Insurance and the Risk Management and Insurance Review published between 2007 and 2010 were selected. The data was analysed by the application of the software AntConc 3.2.1w developed by Anthony (2007).

The results derived from the present study through a corpus-based analysis method revealed a list of keywords which are frequently used in the field of insurance. From the selected top one hundred keywords, the collocates to each keyword were identified. Only collocations that met specified criteria were selected and categorised into lists of 2-, 3- and 4-word clusters.

Many of the keywords and collocations found in this present study could be useful to ESP learners and teachers who are involved in the field of insurance. The findings obtained from this research could be adapted to classroom materials and also references. Many of these collocations would be beneficial to insurance students, teachers, employees, and learners who are interested in the field of insurance.

_____ _____ __ / __ / __

Student's signature Thesis Advisor's signature

ACKNOWLEDGEMENTS

First of all, I would like to express my deep gratitude to my advisor, Dr. Montri Tangpijaikul, for his support, supervision, understanding and valuable advice on this thesis. His knowledge and exceptional experience in corpus study have undeniably shaped my ideas and inspiration. I also must thank the members of my thesis committee, Dr. Prapat Makchuchit and Dr. Issariya Tassana-ngam, for their academic advice and suggestions.

I gratefully acknowledge Assistant Professor Chamaipak Maiklad, Ph.D. for her constructive comments and guidance when I had my thesis proposal presented to the committee.

I would like to convey my sincere gratitude to Ms. Jirawan Iamsooksai and Ms. Siriporn Wongurai for agreeing to be the raters for this thesis. Their expertise in the field of insurance is most appreciated.

To all my classmates in the ESP program, I owe a great deal for their support, comments and encouragement especially in my difficult moments. Thanks to all the wishes, help and friendship along the way of my academic life.

I would like to extend my thanks to Mr. Witsanu Hansuk, who has helped me abundantly in gaining access to volumes of research articles. On top of that, he provided me with encouragement and thoughtful support by sharing his knowledge until I saw this thesis finished. I also thank Mr. Todd Oakes, Mr. Richard Clark and Ms. Katherine Dickinson for helping me with proofreading my thesis.

Last but not least, I wish to express my special thanks to my family for always being supportive and caring. I owe my parents special thanks for giving me strength through encouraging support and love.

Tanu Khamphairoh

October, 2012

TABLE OF CONTENTS

	Page
LIST OF TABLES	iii
LIST OF FIGURES	iv
LIST OF ABBREVIATIONS	v
CHAPTER I INTRODUCTION	1
Rationale of the Study	1
Objectives of the Study	2
Research Questions	3
Significance of the Study	3
Scope of the Study	3
Definition of Terms	4
CHAPTER II REVIEW OF LITERATURE	5
Types of English Vocabulary	5
Keywords	7
The Collocation	8
Research Article	10
A Corpus-Based Approach	11
English in the Insurance Field	13
CHAPTER III METHODOLOGY	16
Research Design	16
Sampling and Data Collection	16
Research Instrument	17
Data Storage and Analysis Procedures	19

TABLE OF CONTENT (CONTINUED)

Page

Verification of the Research Methods and Findings 21

CHAPTER IV RESULTS 23

The List of the Top 100 Keywords 23

Word Frequencies in the IRAC 23

Insurance Keywords in the IRAC 25

Keywords with Meaning Related to Insurance 27

The Collocations of a Keyword 32

CHAPTER V DISCUSSION AND CONCLUSION 48

Grammatical Classification of the Keywords Found in the IRAC 48

Most Frequent Collocations of the Keywords Found in the IRAC 58

Pedagogical Implication 59

Limitations and Recommendations for Further Study 62

Summary of the Thesis 63

REFERENCES 64

APPENDIX 71

BIOGRAPHICAL DATA 105

LIST OF TABLES

Table		Page
4.1	The top 10 keywords in comparison with the BNC Sampler	26
4.2	A rating scale assigned by the researcher and the raters	28
4.3	Inter-rater reliability check with the first rater	30
4.4	Inter-rater reliability check with the second rater	30
4.5	The eliminated thirty-six words previously assigned at step 1	31
4.6	The top 64 keywords with meanings to insurance	31
4.7	Examples of 2-word clusters of the first 10 keywords	33
4.8	Examples of 3-word clusters of the first 10 keywords	38
4.9	Examples of 4-word clusters of the first 10 keywords	43
5.1	Distribution of the 64 keywords across five word classes	50
5.2	Classification of the 58 noun keywords	52
5.3	Vocabulary wordlist distribution of the top 64 keywords	56

LIST OF FIGURES

Figure		Page
2.1	A rating scale for finding technical words	7
3.1	Analysis procedures	19
4.1	Wordlist in the IRAC	24
4.2	Wordlist sorted into lemma list	25
4.3	Modification of the rating scale	28
4.4	A rating scale for finding words with meanings to insurance	28
5.1	Distribution of the 64 keywords across five word classes	49
5.2	Classification of the 58 noun keywords into five areas	51
5.3	Hierarchy of lexical bundles with ‘insurance’ as the node	58

LIST OF ABBREVIATIONS

IRAC = The Insurance Research Article Corpus

AWL = Academic Word List

GSL = General Service List

BNC = British National Corpus

CHAPTER I

INTRODUCTION

Rationale of the Study

It has been well noted that ‘collocation’ plays a major role in language lexical acquisition because of its frequent occurrences (McCarthy, 1990: 12) and *"because the way words combine in collocations is fundamental to all language use"* (Hill, 2000: 53). The term collocation is defined as *“the way in which words co-occur in natural text in statistically significant ways”* (Lewis, 2000: 132). Indeed, knowledge of lexical collocation is necessary for English language learning and teaching. Possessing ability in collocation use can help learners *“think more quickly and communicate more efficiently”* (Hill, 2000: 54). Learning collocation will assist language learners in the way of improving fluency and native-like ability to select appropriate words in context (Shin and Nation, 2008: 340), so it is crucial to learn and possess skills in the use of collocation. In an academic setting, it is important for students to learn high-frequency collocations, because *“high-frequency collocations are part of the lexicon which learners need to acquire”* (Durrant, 2009: 158).

Collocations can be found in all kinds of texts, both spoken and written (Lewis, 2000: 188). But a text from a particular discipline may have different collocations from that of other disciplines. There are many specific terminologies that seem to be ‘domain specific’ (Durrant, 2009: 158). In the past decade, there have been many studies carried out employing a corpus analysis approach with an effort to generate a listing of high-frequency collocations. One in general English can be found in the study of Shin and Nation (2008). Many other corpora focus on specific academic disciplines such as science (Gledhill, 2000a), medicine (Marco, 2000), and engineering (Ward, 2007). There is also a contrastive analysis in cross-disciplinary areas such as history and biology (Cortes, 2004). The most recent study by Durrant

(2009) attempts to study listings of important cross-disciplinary academic collocations, but for some reason the insurance discipline was not included in his focus.

Although many research studies of collocation in academic texts have been carried out throughout the past decade, so far there has been no empirical study of the collocation in the field of insurance. For this reason, this study aims to explore the list of most frequent collocations in insurance research articles appearing in electronic format, so as to offer some useful suggestions and provide references for language learners and teachers in this particular field.

In this study, there are three main reasons for selecting the insurance discipline for a corpus-based study and the research articles in an electronic format in particular. Firstly, the insurance industry is a business activity, whose origins can be traced back to about B.C. 1000 (Prudential Insurance Company of America, 1915). Secondly, insurance has been well-documented and 'Risk Management and Insurance' is one of the academic programs in many universities (Holbrook, 2012). The vocabulary of English texts in the insurance industry is the key to understanding and learning the essential language and terminologies used in the insurance industry for both students and insurance professionals. Lastly, according to Durrant (2009), electronic versions of research articles from online journals can be collected in larger amounts than other types of academic writings such as text books or students' essays, and it is especially suitable to compiling a corpus because transcribing the printed version of text books or students' essays, for example, is time consuming and prone to typing errors.

Objectives of the Study

There are two main purposes as follows:

1. To identify keywords in insurance research articles.

2. To explore the patterns of collocations of the keywords used in insurance research articles.

Research Questions

1. What are the keywords found in insurance research articles?

2. What are the most frequent collocations of the keywords found in insurance research articles?

Significance of the Study

The findings of this study will provide a listing of high frequency collocations in insurance research articles that can be used as a guideline and reference for (a) university students who study and teachers who teach in the insurance field, (b) insurance professionals who learn to communicate English in their workplace, and (c) anyone who is interested in the field of English for the insurance industry. English language learners will be able to use the collocational patterns found in this study to improve their English fluency in the field of insurance.

Scope of the Study

The scope of this study involves corpus size and text selection. The corpus that will be built for this study is the Insurance Research Article Corpus (IRAC). It contains approximately 1 million running words from two journals. However, it is somewhat small when compared to other corpora such as Coxhead's (2000). This is a corpus of 3.5 million running words. This is justifiable since IRAC is a corpus designed for a specific use in the area of insurance only. This study only concentrates on insurance in general, whereas future research could investigate the sub-disciplines of insurance, e.g. property and casualty insurance, health insurance, marine insurance, and miscellaneous insurance.

Definition of Terms

'Collocation' refers to two or more words that often occur together in a text within a space of four words either to the left or right of the selected word (Sinclair, 1991).

'Insurance Research Articles' refers to research articles related to insurance fields. The articles are found in two journals from the international database *Wiley-Blackwell* (www.wiley.com): the *Journal of Risk and Insurance* and the *Risk Management and Insurance Review.*

'Corpus-based Analysis' refers to a systematic approach with a format of computerised assistance that is used to investigate linguistic aspects of language from a large number of authentic texts.

'Keyword' refers to a word that is the key lexical item used in the insurance field. It appears significantly in insurance research articles.

'Node' refers to a word which is selected to find those words which often appear together with it.

'Lemma' refers to the base form of a word

'Lemmatisation' refers to the process of putting all forms of a word together into their lemmas.

CHAPTER II

LITERATURE REVIEW

In this chapter the related literature is presented. It is divided into six sections. The first section discusses the related concepts and previous related studies of the types of English vocabulary. The second section is about keywords. The third is about the concepts and previous related studies of collocation. The fourth presents a review of the research articles. The fifth presents the idea of a corpus-based approach. Finally, English in the field of insurance is discussed in the last section.

Types of English Vocabulary

Based on word frequency, Nation (2001: 11-12) divides words in any text into four groups: (1) high-frequency words, (2) academic vocabulary, (3) technical vocabulary, and (4) low-frequency words.

Firstly, high-frequency words cover 80 per cent of texts. They are basic English words including function words. The attempt to create a list of high-frequency words has been evident in the study of West (1953). The list is called General Service List (GSL), available at http://www.lextutor.ca/freq/lists_download/. It consists of 2,000 word families. Secondly, academic vocabulary is commonly found in academic texts of different genres and fields. It accounts for 8 to 10 per cent of running words in any academic texts. The most widely used academic vocabulary can be seen in Academic Word List (AWL) (Coxhead, 2000), available at http://www.lextutor.ca/freq/lists_ download/. It consists of 570 word families and 3,107 types. For example, there are four word types in the word family of the headword *"abandon"*, including *"abandoned, abandoning, abandonment, and abandons"*. Thirdly, technical vocabulary can be found in specific subject areas and covers up to 5 per cent of texts. Finally, low-frequency words *"occur very infrequently and cover*

only a small proportion of any text" (Nation, 2001: 19). Low-frequency words are the rest of words that do not belong to any of the lists and also do not have meanings related to a particular field, so they are not technical vocabulary.

Previous Related Studies

So far, it appears that the majority of studies on lexical analysis have emphasised the importance of academic vocabulary. The study of using concordancing programmes in learning and teaching academic vocabulary can be found in Thurstun and Candlin (1998) and Sánchez Hernández (2005). Many other corpora focus on using academic vocabulary in specific academic disciplines such as medicine (Chen and Ge, 2007), agriculture (Martínez, Beck, and Panza, 2009), finance (Lia and Qian, 2010), and applied linguistics (Vongpumivitch, Huang, and Chang, 2009). On the contrary, there are only few studies that have focused on technical vocabulary. Furthermore, unlike GSL and AWL, the wordlist for technical vocabulary particularly in the field of insurance has yet to be created. That is because technical words *"differ from subject area to subject area"* (Nation, 2001: 12), and the field of insurance contains many technical words that are used specifically rather than generally.

Because there is no exact number of technical words, there have been attempts to identify lists of technical vocabulary. The four systematic approaches were introduced by Chung and Nation (2004). The first approach is the use of a technical dictionary. With this method, words are considered technical if they occur in the technical dictionary. The second is the use of a corpus-based frequency count by applying a comparative analysis between a general corpus and a specialised corpus. Through the application of statistical formula, words that rank higher in frequency in the specialised field than in other fields are counted technical as stated "*Technical terms are likely to occur only in a specialised field or to occur with a much higher frequency in a specialised field than in a different field or in a variety of other texts"* (ibid.: 259). The third is the use of clues in the text. This method is common in the

anatomy field where technical terms are presented in the forms of definition, typography, and diagrams or illustrations. The last method involves judgment based on expertise in a field and is called 'a four step rating scale' which involves classification of words into four scales by rating their level of technicality. An example from applying a rating scale approach in an anatomy textbook is presented in Figure 2.1. Only those words classified in steps 3 and 4 were counted as technical vocabulary (ibid.).

Step 1: Words with no meaning to anatomy (e.g. *the, is, common, and early*)
Step 2: Words with minimal meaning related to anatomy (e.g. *pairs, structures, and lodges)*
Step 3: Words with related meaning to anatomy (e.g. *chest, neck, abdomen, and cavity)*
Step 4: Words with specific meaning only to anatomy (e.g. *thorax, trachea, and pedicle)*

Figure 2.1 A rating scale for finding technical words (as applied to anatomy text)
Source: Chung and Nation (2004)

Chung and Nation's (ibid.: 262) comparison of the four approaches suggests that the rating scale approach is the most valid compared to the other three approaches. In this study, the concept of the four types of vocabulary (Nation, 2001), the frequency count in identifying keywords, and a rating scale approach (Chung and Nation, op. cit.) are followed.

Keywords

A keyword is defined as *"a word which occurs with unusual frequency in a given text"* (Scott, 1997: 237). With this definition in mind, it is not surprising to find that technical terms are also keywords. But we can only learn that a word appears with an unusually high frequency when we compare its occurrences in the corpus under study against other reference texts. After the compilation of a corpus and

sorting out the lexical items into the wordlist according to frequency, the list is compared against the frequency list of a reference corpus (ibid.).

A word is considered 'key' only when its frequency rank in the corpus under study is high in comparison to its ranks in the reference corpus. A reference corpus that is selected must be sizable (ibid.: 244), either larger or similar size to the one that is being studied (Baker, 2004: 347), or equal to or more than one million words (Scott, op. cit.). The British National Corpus (BNC) Sampler has been widely used as a reference corpus of general English in corpus studies. The corpus comprises written and spoken domains, one million words each. These collections derive from the main British National Corpus (BNC).

The results from a comparison do not guarantee technicality, however, because it may be that not all keywords are technical terms. This is because the procedures of identifying keyness are based on statistical calculation of frequency. In terms of the selection of statistical formula, the two most important are log-likelihood and chi-square. Following the concept of Dunning (1993) that log-likelihood gives the better indication of keyness, many researchers have agreed with this concept and applied this measure to their research instruments (Scott, 1997; Baker, 2004; Durrant, 2008). Therefore, this study uses log-likelihood as it has proved to be superior.

The Collocation

The term 'collocation' was first used by Firth (1957). At the time, he did not provide a definition of this term, but his statement *"you shall know a word by the company it keeps"* (Firth, 1957: 196) has been famously cited. Many works involving collocational studies such as Halliday (1966), Sinclair (1966), and Hoey (1991) have, thereafter, prominently been developed. Recently, O'Dell and McCarthy (2008: 4) defines the meaning of collocation as *"a natural combination of words; it refers to the way English words are closely associated with each other"*.

For Sinclair (1991), collocation is the habitual co-occurrence of two or more words in a text within a space of four words to the left and right of the 'node' which represents an item whose collocations are studied, and those words found within the span are called 'collocates'. It is important to search for collocates in a span of keywords. This is because the chance of finding collocates co-occurring on either side of the node or keyword within a span of four words is greater than finding them adjacent to each other (Durrant, 2008: 85).

The habitual co-occurrence of words has to do with *"the relationship a lexical item has with items that appear with greater than random probability in its (textual) context"* (Hoey, 1991: 7). Put into simple words, the appearance of one word allows the prediction of the appearance of another word rather than an occurrence by chance. This is to say the words would be counted as 'collocates' of the given word only if they are found together more often than their individual frequencies would predict (Jones and Sinclair, 1974).

All in all, the two or more words that often occur together within a space of four words either to the left or right side of the selected word is the concept of collocation which will be used in the present study.

Previous Related Studies on Collocation

There have been several studies carried out recently that aim to provide learners with listings of high-frequency collocations. Shin and Nation (2008) conducted a study to find a list of the highest frequency collocations in spoken English. The study was carried out in the manner of corpus-based analysis, where ten million words in the spoken section of the British National Corpus (BNC) were used as the data source. By analysing the 1,000 most frequently spoken words from the corpus, the researchers extracted 4,698 collocations and for that number a listing of the top 100 collocations was presented. The collocation items in the list were suggested to be learned in an elementary speaking course.

Collocational studies in specific academic disciplines can also be found. Cortes (2004) investigated the presence of lexical bundles in two corpora: research articles and student writings from the history and biology disciplines. The study found a list of 54 and 109 lexical bundles, or clusters of four-word combinations, that were frequently used in history and biology research articles, respectively. A comparative analysis between corpora was employed, and the results showed the relatively low frequency of use of lexical bundles in student writings in both disciplines when they were compared with the frequency of use in published writings. For students to be familiar with and able to use more lexical bundles appropriately, it was suggested that students be exposed to their usage.

Ward (2007) investigated the behaviour of collocations in relation to the lexical technicality in engineering. Two corpora were constructed, and a comparative analysis was conducted among the sub-subjects within the discipline. The results revealed the importance of technical terms and collocations in the form of complex noun phrases commonly found in the study. Ward (2007: 27) also provided insight into the study of collocation in a specific discipline: "*we move from breadth of more general vocabulary knowledge to depth of knowledge of particularly central vocabulary in the student's engineering discipline*".

Research Article

Among other types of academic written texts is the research article (Flowerdew and Peacock, 2001: 96; Bailey, 2006: 3). It is a report that is generally published in an academic journal (Klein-Wohl and Yahav, 2000: 641) or presented in conference proceedings (Rizvi, 2005: 481). Research articles are available both in paper and electronic formats (Harter, 1998: 507). In studies that are corpus-based, an electronic format is more appropriate because it can be adapted directly, whiles the conversion, either by optical scanning or keyboarding, of texts to digital format is needed for paper format (Sinclair, 1991: 14).

The findings presented in research articles are intended to make people in the area of the study agree on their validity and importance (Kushner, 2003: 426), to extend knowledge, and to increase understanding in a specific field of study (Rizvi, 2005: 481). Furthermore, a research article is perceived as a community where scholars can exchange their scholarly communication (Harter, 1998: 507).

Research articles are considered texts of the same genre because they all consist of four main sections (IMRD): Introduction, Method, Results and Discussion (Swales, 1990: 11). In the textual analysis of research articles, a number of studies have been carried out investigating particular sections: *Introduction* (Gledhill, 2000b; Samraj, 2002), *Method* section (Bruce, 2008), *Results* section (Brett, 1994; Bruce, 2009), and *Discussion* section (Holmes, 1997; Hopkins and Dudley-Evans, 1998; Peacock, 2002). It is suggested that in a corpus study, analysing an entire document will allow for the encounter of a wider range of linguistic aspects (Sinclair, 1991: 19).

The research article is recognised as the ideal form of academic writing for students (Hyland, 2008: 47). For the study of English for Academic Purposes (EAP), a corpus of research articles may therefore represent the language academic learners should acquire (Durrant, 2008: 194).

Following the notion of research articles from the above literature, the sampling of this study is research articles from the field of insurance available in electronic format (Harter, 1998: 507), and the whole texts are selected (Sinclair, 1991: 19).

A Corpus-Based Approach

A corpus is generally defined as "*a large collection of written or spoken texts that is used for language research"* (COBUILD, 2001: 338). To be more specific, it is *"a collection of linguistic data, either written texts or a transcription of recorded*

speech, which can be used as a starting-point of linguistic description or as a means of verifying hypotheses about a language" (Crystal, 2003: 112).

Because a corpus is a collection of texts in computer-readable format, a corpus-based approach then brings about the possibility for researchers to carry out their studies in a systematic way with computerised assistance. Sinclair (1991: 4) notes that in implementing a corpus-based approach, researchers can study and reveal various perspectives (e.g. meaning of words, chunks, lemmas and collocations) related to language pattern analyses.

The first step in a corpus study is the corpus creation (Sinclair, 1991: 13), and the two factors that should be taken into consideration in creating a corpus are the corpus size and the representativeness of a corpus (Leech, 1991; Sinclair, 1991; Flowerdew, 1998).

It has been widely discussed what size of a corpus should be studied. The two opposing viewpoints are 'larger' and 'smaller'. Sinclair claimed to support the concept of larger size when he stated that *"a corpus which does not reflect the size and shape of the documents from which it is drawn is in danger of being seen as a collection of fragments where only small-scale patterns are accessible"* (ibid.: 19). The word 'large' as used by Sinclair can be interpreted as 'as large as possible'. In fact, Sinclair (ibid.) specified that *"a corpus needs to contain many millions of words"*. While Sinclair seems to believe in the larger, the better, it is also true that *"a huge corpus does not necessarily 'represent' a language or a variety of a language any better than a smaller corpus"* (Kennedy, 1998: 68). Leech (1991: 10) has noted that the size does not matter most but the representativeness does. For example, in building a large corpus, written language is usually used more than spoken language simply because written data can be collected easier, thus not representing the right proportion between written and spoken language.

The concept of representativeness is concerned with whether a sample represents the population as a whole (Biber, 1993: 244). Kennedy (1998) asserted that *"it is not easy to be confident that a sample of texts can be thoroughly representative of all possible genres or even of a particular genre or subject field or topic.*" However, Biber (op. cit., 246) suggested a 'sampling-frame', the ways a corpus can meet the higher degree of representativeness. It is defined that a sampling is *"an operational definition of the population, an itemized listing of population members from which a representative sample can be chosen"* (ibid. 244). In other words, it is the way to define the boundary of a population and its sample in an explicit figure. To be more concise, Biber (ibid.) gave an example of the way to define the boundary of population and its sample as below:

> The LOB corpus manual (Johansson *et al.,* 1978) is fairly explicit about the sampling frame used: for books, the target population was operationalized as all 1961 publications listed in *The British National Bibliography Cumulated Subject Index, 1960-1964* (which is based on the subject divisions of the Dewey Decimal Classification system), and for periodicals and newspapers, the target population was operationalized as all 1961 publications listed in *Willing's Press Guide* (1961).

Through this literature, the corpus size to be created for this study will fall within the concept of a smaller corpus (Sinclair, 1991) and follow Biber's (1993) sampling frame in terms of creating representativeness.

English in the Insurance Field

It is noticeable that there is a lack of research related to English in the insurance field. The view of English use in insurance is likely to be seen through Business English, as insurance is one of the subject areas in business. Although insurance is a part of business, the language use in this particular field is better regarded as specialised language, rather than business in general (Nelson, 2000).

Speaking of specialised language, technical terminology is one of the obviously distinct features compared to general language (Bowker and Pearson, 2002: 138). According to Nation (2001: 198), a technical word is *"recognizably specific to a particular topic, field or discipline"*. There are three types of users who are involved in using technical terms: namely, professionals or experts who work in the field, teachers and students or semi-experts who may be familiar with and can recognize some of the terms, and non-experts who are not familiar with the terms in the field but have to be involved, such as translators (Bowker and Pearson, 2002: 27-28). Some technical terms associated with insurance that are found particularly in the context of insurance are *'property insurance', 'adverse selection', 'subrogation, and 'coinsurance'.* Some items have general meaning in general language but have specialised meaning in the insurance discipline: for example, *'premium'* means *'great value'* in general contexts, but means *" the amount paid or payable, usually in regular installments, for an insurance policy"* in the insurance context (Collins English Dictionary, 1991: 1227). This is the concept of English in the insurance field which will be used in the present study.

Except for a study of the English language used in the insurance business by Sangvaree (2004), and a few studies on English used in insurance policies conducted in Thailand (Sudsa-Ard, 2001; Chalinrat, 2001) and in The United States of America (Hathaway and Marroso, 2001), there are no empirical studies on the use of English in the insurance field in the aspect of lexical collocation usage. Sangvaree (2004) explored the English skills used and the problematic English usage from the point of view of employees in insurance companies in Thailand. The results indicated that all communicative skills (speaking, listening, writing, and reading) are required in this business, but the degree of regular usage and its necessity are varied due to the nature of the jobs and the positions. However, it appeared that one of the common problems faced by employees is the knowledge of technical terminologies in the insurance field. Sangvaree (2004: 62) concluded:

> The highest severity was the use of an effective business vocabulary including technical in the area of Marine insurance business… The use of English in the area of Marine Insurance Business requires the capacity to use technical business vocabularies.

Even though during the past decade there have been few studies explicitly carried out to analyse the English language used in the insurance field, some studies brought about in-depth perspectives on the topic. Sudsa-Ard's (2001) study of linguistic characteristics of English legal language used in marine cargo insurance policies is a case in point. His study analysed various levels of English language structures including morphology, syntax, and semantics. At the level of morphology he found that there were repetitive patterns of word usage and word choice related to legal language, as well as the specific terms of the field which were specifically found in the insurance business, such as *warranties, insurance, underwriters, and consignee*. However, Sudsa-Ard (ibid.) studied only at the level of morphology rather than the specific aspects of lexical usage and lexical collocations.

Taking the above studies into consideration, the importance of collocations in academic texts is obvious. The previous studies demonstrate the necessity of students' knowledge of a set of collocations in a particular discipline. However, there is no empirical research literature on lexical collocations in the field of insurance. Therefore, there is still much to be investigated, and it is hoped that this study will be valuable to people in the insurance discipline in terms of increasing lexical proficiency.

CHAPTER III

RESEARCH METHODOLOGY

This section presents the design of the present research. The section consists of five main parts: research design, sampling and data collection, research instrument, data storage and analysis procedures, and verification.

Research Design

The present research employs a corpus-based approach in order to explore the list of collocations from keywords found in insurance research articles. First, the IRAC (Insurance Research Articles Corpus) is created. As the researcher was unable to locate any publicly-available corpus of academic English in the field of insurance to use as a basis for identifying keywords, it was necessary to create one to be used for this study. The research articles in the study are selected from journals published between 2007 and 2010. The instrument used for data analysis, based on statistical technique, is *AntConc 3.2.1w* (Anthony 2007).

Sampling and Data Collection

A corpus of research articles in the insurance discipline is compiled. This is called The Insurance Research Articles Corpus (IRAC). In compiling the corpus, the researcher is well aware of the two main concerns: size and representativeness. Taking these matters into consideration, a corpus to be created for this study is a small corpus, and its sampling of the data is aimed to represent English which is used in insurance research articles.

The sample in the corpus is chosen from research articles in two journals from the international database *Wiley-Blackwell* (www.wiley.com): the *Journal of Risk and Insurance* and the *Risk Management and Insurance Review*. The two journals are selected because of their accessibility to institutional subscribers and their international representative databases. Both of them are members of *American Risk and Insurance Association (ARIA)*, which is the premier professional association of insurance scholars. In addition, the *Journal of Risk and Insurance* has an impact factor score of 0.612. The two journals are available both in printed format and electronic format. For this study, the electronic versions of these research articles were downloaded. The corpus contains 155 articles selected from journals published between 2007 and 2010. The reason for this time-frame is that the researcher aims to explore the list of collocations used in research articles which were recently published. The entire text articles are selected in original length. However, there are some parts to be removed, including charts, diagrams, tables, bibliographies, numbers, references, and appendices. The reason why charts, diagrams, tables are removed is that they are uncountable parts. Bibliographies, references, and appendices are also removed because their contents relate to the sources of references, and not the context of insurance. The reason the researcher decided to study the whole texts is that *"a corpus made up of whole documents is open to a wider range of linguistic studies than a collection of short samples"* (Sinclair, 1991: 19).

Research Instrument

The instrument used for analysis in this study is the computer software *AntConc 3.2.1w* (Anthony, 2007). This software is used in corpus study because the software can process data and display it with interpretable results (Anthony, 2004: 7). *AntConc* is free software with a toolkit that is easy to use. The *AntConc 3.2.1w* is the latest version available during the time of conducting this research. There are four main features in this software: *Word list*, *Collocates, Concordance, Keyword list* and *Clusters.* The details of each feature are described below and the applications of these features to this study are presented in the data storage and analysis procedures section.

The *Word list* is used in the process of compiling a listing of vocabulary. The list of vocabulary is generated in alphabetical and frequency order for lexical comparison of texts.

The *Collocates* is a tool used in creating a list of words that co-occur with the selected term. The collocate items are arranged in frequency order, displaying the total number of occurrences on the left and right of the selected term along with the value of a statistical measure. The span of words to the left and right of the selected term can also be manually selected.

The *Concordance* function of *AntConc* is one of the most widely used features in corpus study in the context of linguistic study. This feature is used in finding the typical collocates of given words. The selected words appear with the text to its left- and right- collocates.

The *Keywords* function in AntConc 3.2.1w Tool is used in finding the keywords in IRAC, which are words that appear significantly more frequently in IRAC than they do in a reference corpus. In this case, the reference corpus used in this study is the British National Corpus Sampler containing one million words from written texts. The Keywords function in AntConc Tools also allows the researcher to select the statistical formula, either chi-squared or log-likelihood tests, in calculating the statistical differences of occurrence between the corpora. For this study, the researcher uses the log-likelihood test, as it gives a better indication of keyness (Dunning, 1993).

The *Clusters* is a tool used to identify words that appear to the left or right of a search term or phrase. The cluster can be generated in an ordered list according to frequency.

Data Storage and Analysis Procedures

Figure 3.1 describes the procedures of this study. This section divides the procedures into two steps. The first step involves the compilation of a listing of keywords found in research articles in the field of insurance. The second step is the identification of the collocations of keywords that are obtained from the first step.

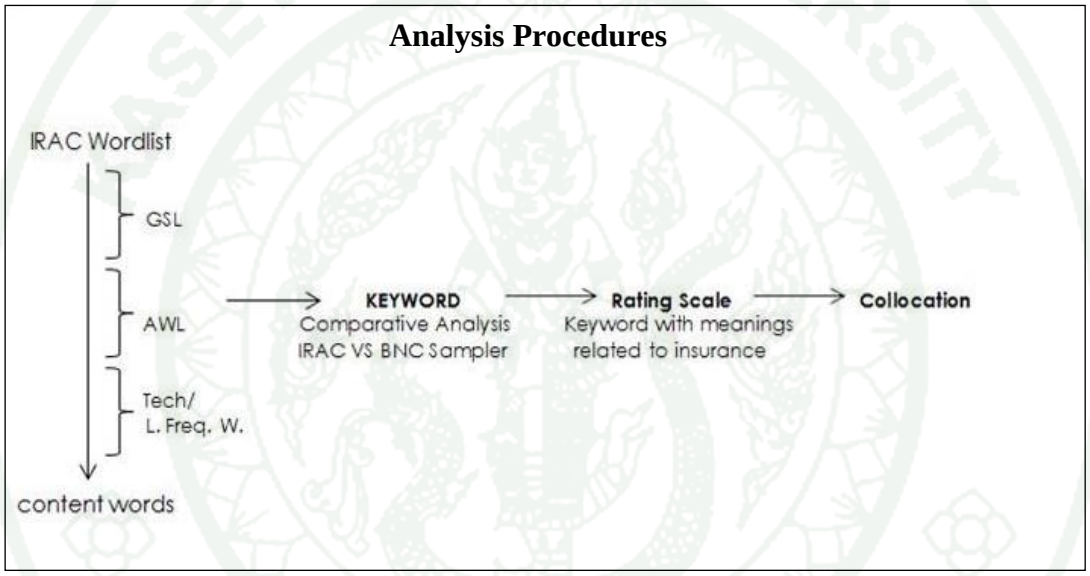

Figure 3.1 Analysis procedures

The First Step: Compiling Corpus and Identifying Keywords

The first step of analysis in this study is the storing of the selected research articles in the IRAC. To standardise all selected research articles, those that contain unreadable components by the computer program such as diagrams, charts, and bibliographies are removed. Once this process is completed the computer software *AntConc 3.2.1w* (Anthony, 2007) with its feature *Wordlist* is employed in order to count the frequency of words found in the entire IRAC. After this process, the word list of the IRAC is identified and stored.

The first research question aims to identify content keywords that are most frequently used in insurance research articles. To answer this question, the 'stop list' feature in Wordlist Tool is used to filter out function words including auxiliary verbs, conjunctions, determiners, prepositions, pronouns, articles, and quantifiers. Thereafter, the *Keywords* feature in *AntConc 3.2.1w* is used to identify a list of words in the IRAC which are statistically significant in their frequency of occurrence. The log-likelihood statistic is used in this procedure. In order to find the list of keywords, the listing of technical words that were stored prior is used to compare its frequency with a larger reference corpus. The reference corpus used in this study is the general English of the British National Corpus Sampler - one million words collected from written texts. This comparative analysis between corpora is based on the assumption that keywords occur significantly more frequently in their discipline but infrequently or not at all in general English. Once this process has finished a listing of the top 100 keywords is created and ready for further analysis. It is worth noting that only those keywords that have a specialised meaning related to the insurance field retrieved by applying *"rating scale approach"* (Chung and Nation, 2004: 252) are selected for further analysis. However, it is noticeable that words categorised in steps 1 and 2 in the rating scale can be considered into the same step as words with general meaning. Additionally, the intention of the rating scale in this study is to identify words that have meanings in the field (words in step 3 and 4) (ibid). Therefore, a simplification from four steps to three steps is made in the present study:

Step 1: Words with general meaning
Step 2: Words with related meaning to the field of insurance
Step 3: Words with specific meaning to the field of insurance

From the above simplified 3 steps, only those that fall under steps 2 and 3 are selected.

The Second Step: Identifying Collocations of Keywords

The second step in the analysis involves identifying collocations of the keywords obtained from the first step. The *Cluster* function of AntConc Tools is used to create listings of the common collocates of each keyword. To produce a careful selection of collocations, the researcher follows two measures. Firstly, the four word span (Jones and Sinclair, 1974) is used, in which the co-occurrence of collocations is limited to occurrences within a four word span. Secondly, the cutoff point of the target collocations is set with a minimum number of four occurrences based on the precedent of four occurrences per one million words for 3-word clusters of keywords (Nelson, 2000). The structural cluster of analysis for this study focuses only on lexical cluster. The clusters which contain function words at the beginning or the end of the cluster are not included in the final list.

Through completing the second step the collocations of the keywords are indicated and compiled as the aim of this study.

Verification of Research Methods and Findings

In this study, the researcher is well aware of the importance of verification of research methods and results. The reliability and validity in this study detail as follows:

The first aspect of concern is the reliability of the instrument. Cohen, Manion, and Morrison (2000: 146) define reliability as *"a synonym for dependability, consistency and replicability over time, over instruments and over groups of respondents"*. The main instrument used in this study is computer software, *AntConc 3.2.1w* (Anthony, 2007). Computers are seen as reliable analysts of text; as Biber (1995: 32) says, *"computers do not become bored or tired"*. Moreover, the authentic texts from research articles used in a corpus represent firm evidence for the study; as

Francis and Sinclair (1994: 191) said *"Corpus data provides us with incontrovertible evidence about how people use language"*.

Cohen, Manion, and Morrison (2000: 133) note on validity that *"in quantitative data validity might be improved through careful sampling, appropriate instrumentation and appropriate statistical treatments of the data"*. In this study, the researcher employs log-likelihood, a statistical test used in *AntConc 3.2.1w* (Anthony, 2007), and in *Wordsmith Tool Version 5.0 c* (Scott, 2009). In WordSmith Help File, Scott (ibid: 160) supports the selection of this test in his software suite because it *"gives a better estimate of keyness, especially when contrasting long texts or a whole genre against your reference corpus"*. Throughout the process of analysis, the researcher also consults with a professor who has knowledge and experience in the use of the software program "*AntConc Tools*". Furthermore, the results from the step of identifying technical vocabulary are obtained by applying a rating scale. The rating scale is perceivably subjective. Therefore, the obtained results are cross checked with two experts in the field of insurance in order to consolidate the accuracy of findings.

CHAPTER IV

RESULTS

In this chapter, the results of the data analysis from the IRAC are presented. The first is the list of the top 100 keywords. The second is the lexical collocations of the keywords.

The List of the Top 100 Keywords

Word Frequencies in the IRAC

From the selected 155 insurance research articles in the IRAC, there were 980,121 running words and 19,476 word types found subject to the analysis with AntConc (Anthony, 2007).

The word frequency list retrieved from the IRAC is shown in a screenshot in Figure 4.1. It is not surprising that function words show the highest frequency rank; similar results are also found in another corpus studies (e.g. Ward, 2009; Sinclair, Jones and Daley, 2004; Scott, 2001). In this corpus the most frequent function word is *the* with over 70,000 occurrences (7.23% of all words), followed by *of* (3.66%), *and* (2.72%), *to* (2.34%) and *in* (2.33%). To filter function words, the function *stoplist* in *AntConc* (Anthony, 2007) is applied. It can be seen that among the top 23 ranked words there are only two content words ranked among the high frequency items: namely, *insurance* (rank 10th) and *risk* (rank 13th). These words are mostly concerned with insurance.

Hits | Total No. of Word Types: 19476 | Total No. of Word Tokens: 980121

Rank	Freq	Word	Lemma Word Form(s)
1	70855	the	
2	35852	of	
3	26667	and	
4	22933	to	
5	22890	in	
6	17969	a	
7	15292	is	
8	12980	that	
9	12725	for	
10	8705	insurance	
11	8299	are	
12	7345	as	
13	7328	risk	
14	6781	with	
15	6486	on	
16	6225	by	
17	6111	this	
18	6008	we	
19	5893	be	
20	4686	s	
21	4313	an	
22	4042	not	
23	3871	from	

Figure 4.1 Wordlist in the IRAC

One last modification needs to be made at this stage. The function *use lemma list file* in *AntConc* is used to assign words into lemmatised lists and count them in the summary of the occurrence of lemma – e.g., the count for ‘ask’ includes ‘ask’, ‘asks’, ‘asking and ‘asked’. The lemma word list can be obtained directly from Laurence Anthony’s Website http://www.antlab.sci.waseda.ac.jp/antconc index.html.

After the exclusion of function words, all data are treated as lowercase. It is notable that words in the entire corpus are treated as lowercase (the option available in *AntConc*) because otherwise, for example, capitalized *Insurance* would be counted as a different type to the normal *insurance*; counting in this way would have affected outcomes significantly. Finally, all words are assigned to lemma word forms. The results are presented in Figure 4.2.

Rank	Freq	Lemma	Lemma Word Form(s)
1	8713	insurance	insurance 8705 insurances 8
2	8369	risk	risk 7328 risks 1041
3	4131	market	market 3102 marketed 4 marketing 53 markets 972
4	4002	insurer	insurer 1389 insurers 2613
5	3123	model	model 2355 modelled 1 models 767
6	3030	loss	loss 1739 losses 1291
7	2862	high	high 1438 higher 1307 highest 117
8	2826	result	result 757 resulted 41 resulting 161 results 1867
9	2768	percent	
10	2698	cost	cost 1479 costs 1219
11	2694	firm	firm 1279 firms 1415
12	2433	increase	increase 910 increased 427 increases 614 increasing 482
13	2429	premium	premium 1388 premiums 1041
14	2367	value	value 1683 valued 38 values 642 valuing 4
15	2144	rate	rate 1299 rated 44 rates 801

Figure 4.2 Wordlist sorted into lemma list

It is notable that the ranks and frequency count in the word list are changed after the exclusion of function words and lemmatisation of the word list. For example, the word *insurance*, which is previously presented in Figure 4.1, includes the frequency of the plural *insurances* (8) and brings its total frequency count to 8713, ranking 1st.

After the modification and application of the software options, the wordlist in the IRAC is ready to be compared with that of the reference corpus, the BNC Sampler written domain, in order to retrieve the list of insurance keywords.

Insurance Keywords in the IRAC

One of the main objectives of the present study is to find the keywords in insurance research articles. Keywords are words that appear with higher frequency rank in the IRAC than in the reference corpus. The computation of the frequency of keywords shows value of keyness of a keyword in form of the statistical basis. In this study, the IRAC is compared to the 1 million words in the written domain of the BNC Sampler corpus to represent general English.

To illustrate the procedure of identifying keywords, an example is provided in Table 4.1, the top 10 keyword list. The words are listed according to their level of keyness in descending order and presented with their frequency and the percentage of occurrences in the entire corpus. The word with higher keyness value is presented with higher rank. The keyness results are calculated by the Log-likelihood statistic, which is an available setting in *AntConc*.

The first column shows the rank of each keyword, followed by the keyword and the keyness score in the third column. The frequency of each word in the IRAC is shown in the fourth column and is followed by its percentage. The next two columns show the frequency in the 1 million words written domain of the BNC Sampler corpus, followed by percentages. The last column presents the difference in percentages of the frequency of occurrences between the two corpora.

Table 4.1 The top 10 keywords in comparison with the BNC Sampler

Rank	Keywords	Keyness	IRAC		BNC Sampler		Difference
			980,121 Tokens		1,183,560 Tokens		(%)
			Frequency	Percentage	Frequency	Percentage	
1	insurance	43488.466	8713	0.889	71	0.006	0.883
2	Risk	41527.172	8369	0.854	91	0.008	0.846
3	insurer	20263.991	4002	0.408	6	0.001	0.407
4	market	17964.193	4131	0.421	468	0.04	0.381
5	model	14677.430	3123	0.319	145	0.012	0.307
6	percent	13946.408	2768	0.282	8	0.001	0.281
7	loss	13784.384	3030	0.309	227	0.019	0.290
8	result	12767.758	2826	0.288	233	0.020	0.268
9	firm	12614.569	2694	0.275	139	0.012	0.263
10	premium	12133.847	2429	0.248	18	0.002	0.246

In the IRAC, the most frequently used word is *insurance* with a keyness score of 43488.466, and appearing 8713 times (0.889% of the entire IRAC). Occurrences of the words *risk* and *insurer* are also extensively found in the IRAC with keyness scores of 41527.172 and 20263.991 and frequencies of 8369 (0.854%) and 4002 (0.408%), respectively. In contrast, the words *insurance*, *risk* and *insurer* occurred much less

frequently in the BNC Sampler, with frequencies of 71 (0.006%), 91 (0.008%) and 6 (0.001%), respectively. This clearly shows that these words are notably common in insurance research articles and not frequently used in general written English. It is obvious that a keyword analysis clearly highlights the terminology related to the insurance discipline. Other words such as *market*, *model*, *loss* and *premium* also have high keyness in this corpus and indicate what the genre of the texts in the corpus is about.

Because keyness is statistically calculated by the comparison of the frequency of the word between two corpora, the word with comparatively higher frequency receives the higher keyness score. A word with higher frequency in the target corpus is not necessarily ranked higher if its frequency is less outstanding after comparison with the reference corpus. This confirms the notion of keyword given by Scott (1997) and explains why in Table 4.1 the word *market* (ranked 4th) is preceded by *insurer* (ranked 3rd), and *loss* (ranked 7th) is proceeded by *percent* (ranked 6th).

Keywords with Meaning Related to Insurance

As the purpose of this study is to focus only on those keywords with semantic meaning related to the insurance field, keywords with no meanings to insurance are then excluded. A rating scale is used to consolidate the accuracy of the results by the researcher and the two raters. It is important to note that in this study there is a simplification of the application of the rating scale. It is clear that in this process, the objective focuses on the identification of words with related and/or specific meaning to insurance (step 3 & 4). Furthermore, only words in steps 3 and 4 are to be selected for further study. Therefore, in the view of the researcher, there is no necessity to classify words with no meaning and/or little meaning (step 1 & 2) into two different steps. For this reason, the four-step rating scale is modified to three steps described in Figure 4.3, and the three-step rating scale used in this study is presented in Figure 4.4 below.

Steps			
Chung & Nation	1-2	3	4
IRAC	1	2	3

Figure 4.3 Modification of the rating scale.

. The three-step rating scale used in this study:

Step 1: Words with no and/minimal meaning to insurance
Step 2: Words with related meaning to insurance
Step 3: Words with specific meaning to insurance

Figure 4.4 A rating scale for finding words with meanings to insurance

The data in Table 4.2 shows all 100 keywords assigned to the three steps by the researcher and the two raters. For most of the keywords the raters agreed with the researcher on the steps assignment. Those viewed differently are marked in bold.

Table 4.2 A rating scale assigned by the researcher and the raters

No.	Keyword	Rated by R.	Rated by 1st rater	Rated by 2nd rater	No.	Keyword	Rated by R.	Rated by 1st rater	Rated by 2nd rater
1	insurance	3	3	3	51	significant	1	1	1
2	risk	2	2	2	52	assume	1	1	1
3	insurer	3	3	3	53	reinsurance	3	3	3
4	market	2	2	2	54	function	2	2	2
5	model	2	2	2	55	analysis	2	2	2
6	percent	1	1	1	56	utility	2	2	2
7	loss	2	2	2	57	optimal	2	2	2
8	result	1	1	1	58	table	1	1	1
9	firm	2	2	2	59	ratio	2	2	2
10	**premium**	**2**	2	**3**	60	plan	2	2	2
11	cost	2	2	2	61	demand	2	2	2
12	high	1	1	1	62	distribution	2	2	2
13	increase	1	1	1	63	purchase	2	2	2
14	asset	2	2	2	64	wealth	2	2	2

Table 4.2 (Continued)

No.	Keyword	Rated by R.	Rated by 1st rater	Rated by 2nd rater	No.	Keyword	Rated by R.	Rated by 1st rater	Rated by 2nd rater
15	rate	2	2	2	65	information	2	2	2
16	price	2	2	2	66	factor	2	2	2
17	value	2	2	2	67	lower	1	1	1
18	coverage	2	2	2	68	hedge	2	2	2
19	effect	1	1	1	69	management	2	2	2
20	**variable**	**2**	2	**1**	70	difference	1	1	1
21	policy	2	2	2	71	investment	2	2	2
22	claim	2	2	2	72	annuity	2	2	2
23	use	1	1	1	73	benefit	2	2	2
24	health	2	2	2	74	stock	2	2	2
25	individual	2	2	2	75	term	2	2	2
26	**expect**	**1**	**2**	1	76	different	1	1	1
27	level	2	2	2	77	income	2	2	2
28	state	1	1	1	78	equation	1	1	1
29	contract	2	2	2	79	bond	2	2	2
30	company	2	2	2	80	property	2	2	2
31	capital	2	2	2	81	average	2	2	2
32	data	1	1	1	82	pay	2	2	2
33	provide	1	1	1	83	consumer	2	2	2
34	also	1	1	1	84	**equilibrium**	**2**	2	**1**
35	**measure**	**2**	2	**1**	85	test	1	1	1
36	financial	2	2	2	86	consider	1	1	1
37	large	1	1	1	87	suggest	1	1	1
38	article	1	1	1	88	selection	2	2	2
39	year	2	2	2	89	section	1	1	1
40	case	2	2	2	90	**worker**	**1**	**2**	1
41	period	2	2	2	91	reduce	1	1	1
42	study	1	1	1	92	example	1	1	1
43	probability	2	2	2	93	only	1	1	1
44	low	1	1	1	94	rating	2	2	2
45	**estimate**	**1**	**2**	1	95	group	2	2	2
46	life	2	2	2	96	private	2	2	2
47	return	2	2	2	97	reinsurer	3	3	3
48	base	1	1	1	98	relate	1	1	1
49	sample	1	1	1	99	show	1	1	1
50	liability	2	2	2	100	positive	1	1	1

To check the reliability of the assignment of the 100 words in Table 4.2, an inter-rater reliability check with the raters is implemented and the accuracy score is presented in Table 4.3 (score between the researcher and the first rater) and Table 4.4 (score between the researcher and the second rater).

Table 4.3 Inter-rater reliability check with the first rater

Steps rated by the 1st rater	Steps rated by the researcher			Total words assigned by the rater
	1	2	3	
1	**33**			33
2	3	**60**		63
3			**4**	4
Total words assigned by the researcher	36	60	4	Accuracy score **(33+60+4) = 97%**

As the data shows in Table 4.3, the first rater agreed with the researcher on the assignment of words at step 3. At step 1, the first rater agreed with the researcher on 33 out of 36 words, with the other 3 including *expect, estimate* and *worker* are assigned to step 2. At step 2, there are 60 words with agreement between the first rater and the researcher. Thus, the total accuracy score of the reliability check with the first rater is (33+60+4) = 97%.

Table 4.4 Inter-rater reliability check with the second rater

Steps rated by the 2nd rater	Steps rated by the researcher			Total words assigned by the rater
	1	2	3	
1	**36**	3		39
2		**56**		56
3		1	**4**	5
Total words assigned by the researcher	36	60	4	Accuracy score **(36+56+4) = 96%**

It can be seen in Table 4.4 that the rater agreed with the researcher on assigning 36 words at step 1, 56 words at steps 2, and 4 words at step 3. There are 3 words the researcher assigned at step 2 including *variable, measure* and *equilibrium* which the rater assigned at step 1. At step 3, besides *insurance, insurer, reinsurance* and *reinsurer*, the rater also assigned *premium*, whereas it is assigned at step 2 by the

researcher. The total accuracy score of the reliability check with the second rater is (36+56+4) = 96%.

After implementation of the reliability check with the two experienced experts in the insurance filed, 36 keywords which in the rating scale were assigned at step 1 are eliminated, as displayed in Table 4.5.

Table 4.5 The eliminated thirty-six words previously assigned at step 1

1	high	11	expect	21	sample	31	reduce
2	percent	12	provide	22	assume	32	worker
3	result	13	only	23	lower	33	positive
4	increase	14	study	24	difference	34	relate
5	also	15	low	25	show	35	different
6	use	16	table	26	section	36	example
7	state	17	article	27	test		
8	effect	18	base	28	consider		
9	data	19	significant	29	equation		
10	large	20	estimate	30	suggest		

After fulfilling the criteria of word selection by applying a rating scale plus the inter-rater reliability check, Table 4.6 presents the ultimate list of the final 64 keywords with meanings related to the field of insurance.

Table 4.6 The top 64 keywords with meanings to insurance

1	insurance	17	claim	33	reinsurance	49	annuity
2	risk	18	health	34	function	50	benefit
3	insurer	19	individual	35	analysis	51	stock
4	market	20	level	36	utility	52	term
5	model	21	contract	37	optimal	53	income
6	loss	22	company	38	ratio	54	bond
7	firm	23	capital	39	plan	55	property
8	premium	24	measure	40	demand	56	average
9	cost	25	financial	41	distribution	57	pay

Table 4.6 (Continued)

10	asset	26	year	42	purchase	58	consumer
11	rate	27	case	43	wealth	59	equilibrium
12	price	28	period	44	information	60	selection
13	value	29	probability	45	factor	61	rating
14	coverage	30	life	46	hedge	62	group
15	variable	31	return	47	management	63	private
16	policy	32	liability	48	investment	64	reinsurer

The Collocations of a Keyword

In this section the process of identifying collocations is carried out to answer the second research question – *what are the most frequent collocations of the keywords found in insurance research articles?* The results are shown the presentation of collocations. The top 64 keywords as presented earlier in Table 4.6 are chosen for further analysis in order to see their collocational patterns in insurance research articles. These keywords are reasonably chosen. It is not easy to investigate the large numbers of keywords found in a corpus study (Nelson, 2000), so priority is given to the keywords that are most frequent.

To identify which lexical items typically collocate to the left or the right of the selected keywords, the *Cluster* function in *AntConc* is used. The twenty-five most frequent clusters with at least 4 occurrences are selected. For 2-word clusters, only lexical collocates are included. For 2-4 word clusters, the clusters including function words at the first and/or last position of the cluster are excluded on the ground that this study focuses only on lexical collocations. The comprehensive list of the 2-, 3-, and 4-word clusters keyword collocations is shown in Appendix. Shown in Table 4.7 is the list of the first 10 keywords with their top five 2-word clusters to the left and right of the node.

Table 4.7 Examples of 2-word clusters of the first 10 keywords

No.	Keywords	2 –word clusters		
1	insurance	life	insurance	
		health	insurance	
		purchase	insurance	
		liability	insurance	
		private	insurance	
			insurance	market
			insurance	company
			insurance	coverage
			insurance	industry
			insurance	demand
2	risk	low	risk	
		high	risk	
		mortality	risk	
		asset	risk	
		background	risk	
			risk	management
			risk	aversion
			risk	free
			risk	type
			risk	minimization
3	insurer	life	insurer	
		stock	insurer	
		international	insurer	
		liability	insurer	
		primary	insurer	
			insurer	financial
			insurer	participation
			insurer	performance
			insurer	stock
			insurer	size
4	market	insurance	market	
		capital	market	
		stock	market	
		reinsurance	market	
		financial	market	
			market	share
			market	value
			market	concentration
			market	based
			market	condition

Table 4.7 (Continued)

No.	Keywords	2 –word clusters		
		feedback	model	
		basic	model	
		pricing	model	
		selection	model	
5	model	simple	model	
			model	framework
			model	parameters
			model	includes
			model	risk
			model	using
		expected	loss	
		earnings	loss	
		aggregate	loss	
		developed	loss	
6	loss	potential	loss	
			loss	coverage
			loss	ratios
			loss	distribution
			loss	experience
			loss	probability
		insurance	firm	
		new	firm	
		individual	firm	
		financial	firm	
7	firm	breached	firm	
			firm	size
			firm	specific
			firm	characteristics
			firm	performance
			firm	level
		insurance	premium	
		pure	premium	
		risk	premium	
		fair	premium	
8	premium	average	premium	
			premium	growth
			premium	income
			premium	rate
			premium	subsidies
			premium	paid

Table 4.7 (Continued)

No.	Keywords	2 –word clusters		
		total	cost	
		high	cost	
		low	cost	
		hedging	cost	
9	cost	average	cost	
			cost	sharing
			cost	function
			cost	efficiency
			cost	growth
			cost	containment
		risky	asset	
		optimal	asset	
		illiquid	asset	
		initial	asset	
10	asset	liquid	asset	
			asset	allocation
			asset	risk
			asset	reconstitution
			asset	class
			asset	return

'Insurance'

Insurance market is the most frequent of the 2-word clusters with collocates to the right and *life insurance* with collocates to the left of the keyword *insurance* in the IRAC. *Insurance* displays strong meanings when describing the types of insurance products, e.g. *life insurance*, *health insurance*, *liability insurance* and *private insurance*. Combined with collocates to the right, it is found the formation of words as compound nouns, e.g. *insurance market*, *insurance company*, *insurance coverage*, *insurance industry* and *insurance demand*. This is why it is good to learn these as one lexical item.

‘Risk’

Learners need to know that risk conveys negative meaning, and that it is something to be avoided, as can be seen in *risk aversion*, or managed with care, as in *risk management*. If not it could be fatal, as seen in *mortality risk*. Risk comes in degree and appears with the adjectives *high* and *low*.

‘Insurer’

Collocates to the left of *insurer* describe the type of insurer – *life insurer*, *stock insurer* and *international insurer*. Insurer is the principal institute in the insurance industry and engages in activities in the insurance business. This engagement is called *insurer participation*, and insurers’ performance is known as *insurer performance* for clients.

‘Market’

In the IRAC, *market* is particularly concerned with the types of market in the field of insurance, as can be seen in *insurance market*, *capital market* and *stock market*. The word ‘market’ is used mainly as a noun. When it functions as an adjective, it refers to the prices of things that are not fixed. For example, *market value* refers to the value determined by the demand and supply of the economy system.

‘Model’

Insurance is an industry that involves risk and possibility of losses, so there are standardised models in predicting future events, as seen in *feedback model*, *basic model*, *pricing model*, *model framework* and *model parameters*.

‘Loss’

Loss is a direct result of risk. In this study, *loss* is found with meanings which refer to the types of losses, such as *expected loss*, *earnings loss* and *aggregate loss*. We purchase an insurance policy so as to have our losses covered (*loss coverage*) and distributed (*loss distribution*) by insurers.

‘Firm’

The word *firm* mostly collocates with adjectives that suggest its qualities, such as *new firm*, *individual firm*, and *financial firm*. Firm is also referred to in terms of its size, characteristics, performance and level, as in *firm size*, *firm characteristics*, *firm performance* and *firm level*.

‘Premium’

The word *premium* is found to collocate with a wide range of words related to the terms of an insurance policy, as in *insurance premium*, *pure premium* and *premium rate*. It also collocates with words involving financial investment, as in *premium growth* and *premium income*.

‘Cost’

It can be seen in this study that the word *cost* is likely to be discussed as something that needs to be managed, as in *cost sharing* and *cost efficiency*. Cost is often shown in degree adjectives such as ‘*high*’ and ‘*low*’, and adjectives suggesting the amount of something, as in *total cost* and *average cost*.

‘Asset’

The word *asset* is found to collocate frequently with words whose meanings are related to economy and finance, e.g. *asset allocation* and *asset reconstitution*. The combination of *asset* with its collocates also shows the types of assets, e.g. *risky asset*, *optimal asset* and *illiquid asset*.

Table 4.8 below presents the 3-word clusters of the first 10 keywords found in the IRAC. Only the top 5 clusters of each keyword are shown. The full list can be found in Appendix , in which it is presented together with the 2- and 4-word clusters.

Table 4.8 Examples of 3-word clusters of the first 10 keywords

No.	Keywords	3 –word clusters		
		purchase of	insurance	
		cost of	insurance	
1	insurance	lines of	insurance	
		association of	insurance	
		life	insurance	industry
			risk	and insurance
		aggregate mortality	risk	
2	risk		risk	minimization hedging
		relative	risk	aversion
			risk	free rate
		international	insurer	participation
			insurer	stock returns
3	insurer		insurer	investment policy
		life	insurer	stock
			insurer	solvency situation
		life insurance	market	
		cat bond	market	
4	market	stock	market	response
		auto insurance	market	
		capital	market	line
			model	of insurance
		asset pricing	model	
5	model	adverse selection	model	
		jump diffusion	model	
			model	of participating

Table 4.8 (Continued)

No.	Keywords	3 –word clusters		
		probability of	loss	
		excess of	loss	
6	loss	industry	loss	warranties
			loss	reserve estimation
		changes in	loss	
			firm	specific factors
			firm	specific characteristics
7	firm	relation between	firm	
			firm	fixed effects
		associated with	firm	
		stock risk	premium	
		pure	premium	rate
8	premium	actuarially fair	premium	
		fair	premium	rate
		stop loss	premium	
			cost	of capital
			cost	of insurance
9	cost		cost	of equity
			cost	per ship
			cost	of public
		optimal	asset	allocation
			asset	allocation problem
10	asset		asset	risk taking
		risky	asset	reconstitution
		capital	asset	pricing

The number of occurrences of 3-word clusters is considerably less than the 2-word clusters as presented in Table 4.7. It would seem that the longer clusters appear lower in frequency than the shorter ones. However, it can also be seen from the table above that when keywords are combined in a bigger cluster, they seem to take on meanings more specific to the insurance field. For example, the word *market* appears with high frequency in 2-word clusters, especially with *insurance* as can be seen in *insurance market*. However, meanings become clearer with a focus on the types of insurance market in the IRAC, as in *life insurance market* and *auto insurance market*.

'Insurance'

In the IRAC, there appears a typical pattern of a noun + preposition (of) + the word *insurance*, as in *purchase of insurance, cost of insurance, lines of insurance* and *association of insurance*. In the insurance business there are four major *lines of insurance*, including property, casualty, life, and health and disability. Insurance also refers to a group of people who are directly involved in the insurance business, e.g. *association of insurance* and *life insurance industry*. When the meanings refer to an insurance activity and pricing, they are found in the clusters *purchase of insurance* and *cost of insurance*.

'Risk'

This word is commonly found in the clusters *risk and insurance* which indicates that insurance is the branch of risk concerned with transferring risk of loss to insurers. Two other clusters are related to strategies in risk management, e.g. the prediction of the mortality rate as in *aggregate mortality risk*, and the risk of minimum return in investment as in *risk free rate*. The last two clusters seem to have direct meanings to risk measures in financial trading strategy, as in *risk minimization hedging*, and measures that concern the uncertainty of gaining against losing in investment, as in *relative risk aversion*.

'Insurer'

The collocates of *insurer* indicate a focus on insurance activities, especially in terms of money and investment, e.g. *insurer stock returns*, *insurer investment policy* and *life insurer stock*. To ensure that an insurer meets financial obligation to cover liability, the use of the word *insurer* is found in the cluster *insurer solvency situation*.

‘Market’

The word *market* is commonly used to indicate the types of insurance markets, such as *life insurance market*, *cat bond market*, and *auto insurance market*. It is worth noting that *cat* is the abbreviation for ‘Catastrophe bonds’, which is one of the debt instruments in risk protection. The word *market* is also found in the cluster *stock market response*, which refers to the reaction of the market to the changes in a stock. It is also used in a cluster that has direct meanings to financial markets: *capital market line*. The capital market is where people or organisations trade in securities such as bonds and stocks. So, *capital market line* is the line of expected rates of return for the change in the level of risk.

‘Model’

Frequently used 2-word clusters of the word *model* generally have meanings focused on methods of measurement; the 3-word clusters present types of such methods as in *asset pricing model*, *adverse selection model* and *jump diffusion model*.

‘Loss’

Loss is perceivably unpredictable so insurers need to consider risks that could contribute to losses, as seen in *possibility of loss*. Because of this uncertainty, the insurer and reinsurer need to be focused on the adequacy of monetary arrangement in case there is a loss payment, e.g. *loss reserve estimation*. When a loss is over the limit to be covered by the primary insurer - *excess of loss* - it is then shifted to the reinsurer. Among many products of reinsurance, *industry loss warranties* is found in the IRAC. This product provides coverage of catastrophe loss to an industry.

‘Firm’

The word *firm* collocates with words that refer to characteristics, as in *firm specific characteristics*. *Firm* is found in clusters whose meanings refer to relations, such as *relation between firm* and *associated with firm*.

‘Premium’

In the IRAC *premium* collocates with words that refer to the calculation of premium rate, e.g. *pure premium rate, actuarially fair premium* and *fair premium rate*. For example, *pure premium rate* is the estimate of the future costs of losses arising from claims as indicated in historical experience.

‘Cost’

In the IRAC there is also found the collocational pattern of the word *cost* + preposition (of) + a noun. For example, *cost of capital* refers to the opportunity cost of an investment; *cost of insurance* literally refers to the rate of premium buyers need to pay; and *cost of equity* is the expected annual rate of return from investment in stocks and shares.

‘Asset’

Collocates with words referring to economy and finance as found in the results of 2-word clusters, e.g. *optimal asset allocation, risky asset reconstitution* and *capital asset pricing*. The cluster *optimal asset allocation* is the combination of *optimal asset* and *asset allocation*. The cluster *optimal asset allocation* has meanings concerned with the selection of the best way to allocate assets in an investment portfolio. A similar combination is also found in *risky asset reconstitution*¸ in which the cluster *risky asset* is combined with *asset reconstitution*.

The following Table 4.9 presents the list of examples of 4-word clusters of the first 10 keywords from the IRAC. Only the top 5 clusters of each keyword are shown. For a fuller list please see Appendix.

Table 4.9 Examples of 4-word clusters of the first 10 keywords

No.	Keywords	4 –word clusters		
1	insurance	risk management and	insurance	
		demand for life	insurance	
		association of	insurance	commissioners
		life	insurance	and annuities
		demand for health	insurance	
2	risk		risk	management and insurance
		coefficient of	risk	aversion
			risk	minimization hedging using
		constant relative	risk	aversion
		high general	risk	policyholders
3	insurer	life	insurer	stock returns
			insurer	sided adverse selection
		volatility of an	insurer	
		firm and the	insurer	
		number of	insurer	insolvencies
4	market	medical malpractice insurance	market	
		equilibrium in a	market	
			market	for medical malpractice
			market	value of equity
		percent of the	market	
5	model	capital asset pricing	model	
			model	of insurance demand
			model	of employer behavior
			model	for the insurance
		results of the	model	
6	loss	size of the	loss	
		changes in the	loss	
		increase in	loss	coverage
		improvement in	loss	coverage
		premium to expected	loss	
7	firm	value of the	firm	
		take over the	firm	
		size of the	firm	
			firm	and the insurer
		worker	firm	size exclusion
8	premium		premium	to expected loss
		effect of	premium	growth
		equity	premium	of stocks
		percent of the	premium	

Table 4.9 (Continued)

No.	Keywords	4 –word clusters		
9	cost		cost	of public funds
		amount of	cost	sharing
			cost	of equity capital
			cost	of risk management
			cost	and revenue efficiency
10	asset	capital	asset	pricing model
			asset	allocation and liquidation
		solutions to the	asset	
		optimal	asset	allocation given
		case of an	asset	

The 4-word clusters shown in Table 4.9 are similar to the 3-word clusters shown in Table 4.8. The occurrences of the cluster are much less frequent when the word combination is longer, and the meanings are more specific to the insurance field, e.g. *medical malpractice insurance market*, *risk management and insurance*, *asset allocation and liquidation* and *model of insurance demand*.

'Insurance'

The most frequently used 4-word cluster in the IRAC is *risk management and insurance*. This could be due to the fact that *'Risk Management and Insurance Review'*, which is the name of the selected journal for this study, is frequently mentioned. There is also the common pattern of demand + preposition (for) + type of insurance, for example, *demand for life insurance* and *demand for health insurance*.

'Risk'

The word *risk* is frequently used in the 4-word cluster *risk management and insurance* as it refers to subject, association and society. More specific to the function of measurement is risk calculation. It is found in the cluster *coefficient of risk aversion* which is the method of measuring the variables of the degree of taking risks.

Another example can be seen in the cluster *constant relative risk aversion*, a measure of risk aversion which is used to measure the degree of risk taking.

'Insurer'

The word *insurer* collocates with words that display negative meanings, e.g. *number of insurer insolvencies*, which could lead to the idea that an insurance company is facing a lack of financial liquidity. It is also found that *insurer* collocates with positive clusters, as in *life insurer stock returns* denoting the earnings from investment.

'Market'

The word *market* has meanings related to insurance products for professional liability in the medical field, for example, *medical malpractice insurance market* is one of the types of insurance products available for professionals such as doctors who might be exposed to the risk of medical malpractice lawsuits. It is also associated with finance and money as found in *equilibrium in a market*, *market value of equity*, and *percent of the market*.

'Model'

It refers to the types or names of the models, e.g. *capital asset pricing model*¸ *model of insurance demand* and *model of employer behavior*. In the IRAC, it is found that the use of the cluster *model of employer behavior* in the meanings refers to the customisation of insurance coverage options that employers offer to their employees. It is also found in *capital asset pricing model*, with a meaning related to a measure of the liquidity of the asset and the expected return of a security in the market portfolio.

‘Loss’

It is found that the word *loss* combines with words to form a compound noun in the common pattern: positive noun + preposition (in) + loss coverage, e.g. *increase in loss coverage* and *improvement in loss coverage*. *Loss* with its collocate *coverage* in this pattern seems to express positive connotation as something that can be increased or improved.

‘Firm’

The word *firm* is accompanied by words that suggest characteristics of the firm such as *value of the firm* and *size of the firm*. An activity referring to the firm can be seen in *take over the firm*.

‘Premium’

There are only four clusters of the word *premium* found in the IRAC. They display meanings related to finance and investment, as in *effect of premium growth*, *equity premium of stocks*, and *percent of the premium*. These clusters are clearly concerned with the growth of equity gain from the investment in a stock, measured in percentage of expected return of the premium.

‘Cost’

The clusters show the same pattern found in the 3-word clusters of the word *cost*: cost + preposition (of) + noun, e.g. *cost of public funds*, *cost of equity capital* and *cost of risk management*. The first two clusters have meanings referring to the investment as in *cost of public funds* and *cost of equity capital*, and the latter is about the insurance itself as in *cost of risk management*.

‘Asset’

It is found the 4-word clusters *capital asset pricing model* with the highest frequency for the keywords *asset* and *model*. It is noted that both terms are part of the same cluster. For 4-word clusters, *asset* is still concerned with financial aspects, as seen in *asset allocation and liquidation*.

This chapter has presented the results of this study. Some examples have been presented here. The first part shows the results related to the findings of the list of the most frequently used keywords with meanings related to insurance found in the IRAC. The second part reveals the collocational patterns of the selected keywords. These results are analysed and discussed in the next chapter along with limitations and recommendations for further study, and summary of the study.

CHAPTER V

DISCUSSION AND CONCLUSION

This chapter consists of five sections: namely, grammatical classification of the keywords found in the IRAC, most frequent collocations of the keywords found in IRAC, pedagogical implication, limitations and recommendations for further study, and summary of the thesis.

In this chapter, the discussion of the research results is presented to answer the following research questions:

1. What are the keywords found in insurance research articles?
2. What are the most frequent collocations of the keywords found in insurance research articles?

The presentation of the results is begun with discussion of the keywords of the IRAC and followed by discussion of the most frequent collocations of the selected keywords.

Grammatical Classification of the Keywords Found in the IRAC

In terms of the first research question, the quantitative results have provided a list of the first 64 keywords with semantic meanings related to insurance which are frequently used in the selected insurance research articles of the two journals. The result of the keywords in this study reflects the characteristics of the insurance industry. It is not surprising to see the word *insurance* ranked at the top of the list and followed by *risk* and *insurer*. As can be seen, extracting keywords from the corpus makes it possible to give an idea what the corpus in the study is about.

The 64 lexical keywords can be distributed into five word classes: nouns, verbs, adjectives, nouns/adjectives and nouns/verbs.

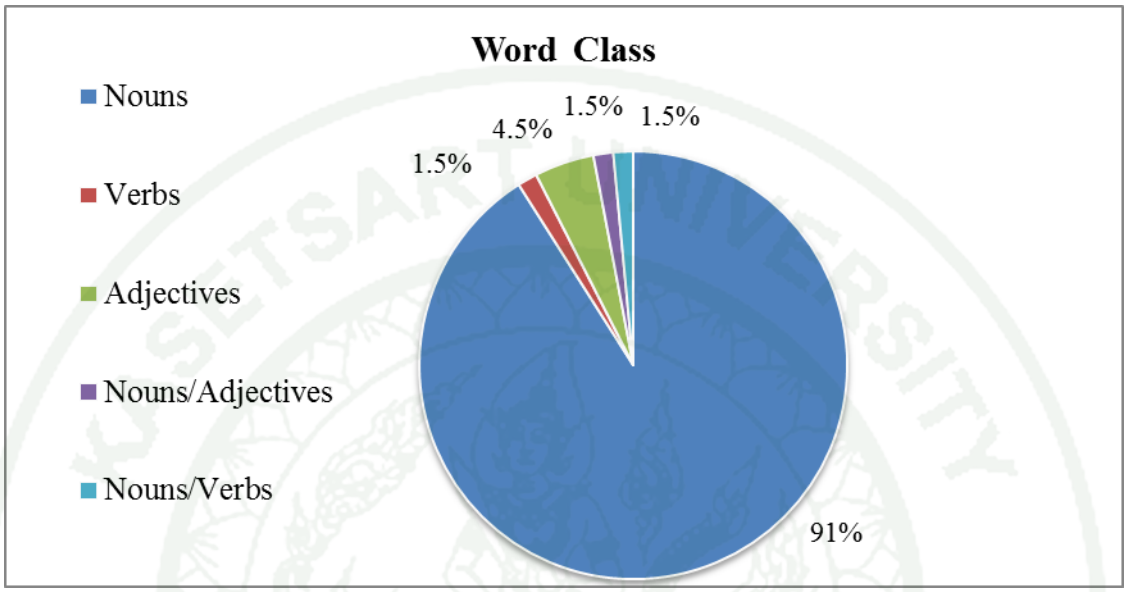

Figure 5.1 Distribution of the 64 keywords across five word classes

The above diagram reveals that among the top 64 keywords 91% of them are nouns which make up the major proportion, followed by adjectives which account for only 4.5%. Verbs, nouns/adjectives and nouns/verbs equally make up 1.5% each. This is not surprising as the most important terms in insurance have meanings related to money and its value (e.g. *cost*, *price*, *value*, *property*, *fund*, *asset*, *equity*, *liability*, *wealth*), as well as with measures or regulations (e.g. *contract*, *measure*, *claim*, *average*). Meanwhile, words such as *purchase* can be both a noun and a verb, as in clusters such as *purchase full coverage* where *purchase* functions as a verb, and *mandatory purchase* where *purchase* functions as a noun. *Average* can be both a noun and an adjective. In the cluster *historical average*, *average* functions as a noun. In the cluster *average income*, *average* functions as an adjective. The full list of word class distribution is presented in Table 5.1.

Table 5.1 Distribution of the 64 keywords across five word classes

Word Classes	Keywords
N (58)	insurance risk insurer market model loss firm premium cost asset rate price value coverage variable policy claim health individual level contract company capital measure year case period probability life liability reinsurance function analysis utility ratio plan demand distribution wealth information factor hedge management investment annuity benefit stock term income bond property consumer equilibrium selection rating reinsurer return group
V (1)	pay
Adj (3)	financial optimal private
N/Adj (1)	average
N/V (1)	purchase

From word class distribution, the 64 keywords can be categorised further into different areas according to their associated meanings. It is important to note that the categorisation of these items may not be ideal or absolute, since one word may fit into two or more categories and others may have a different view or perspective in classifying the meanings of words. The semantic distribution of each word class is presented as follows:

Nouns:

Among the 58 nouns, it can be seen that keywords referring to people, industries and institutes include *insurer*, *firm*, *individual*, *company*, *consumer*, and *reinsurer*. From these six noun keywords, only two seem to have specific references to people in the field of insurance which are *insurer* and *reinsurer*. Keywords related to economy and finance in general include words such as *demand*, *value*, *price*, *variable*, *ratio*, *probability*, *wealth*, *rate*, *asset*, *hedge*, and *stock*. The findings of insurance terms alongside general terms in economics and finance are not at all surprising, as insurance is one kind of financial activity.

The next category contains keywords referring to activities such as *insurance*, *risk*, *reinsurance*, *market*, *analysis*, *distribution*, and *management*. These items reflect the nature of running business activities in insurance field. Products and contracts that are essential to insurance activities include *health*, *life*, *coverage*, *policy*, *liability*, *property*, *premium*, and *claim*. The measures and conditions that have a relationship to insurance products and contracts include *model*, *level*, *measure*, *year*, *period*, *function*, and *factor*.

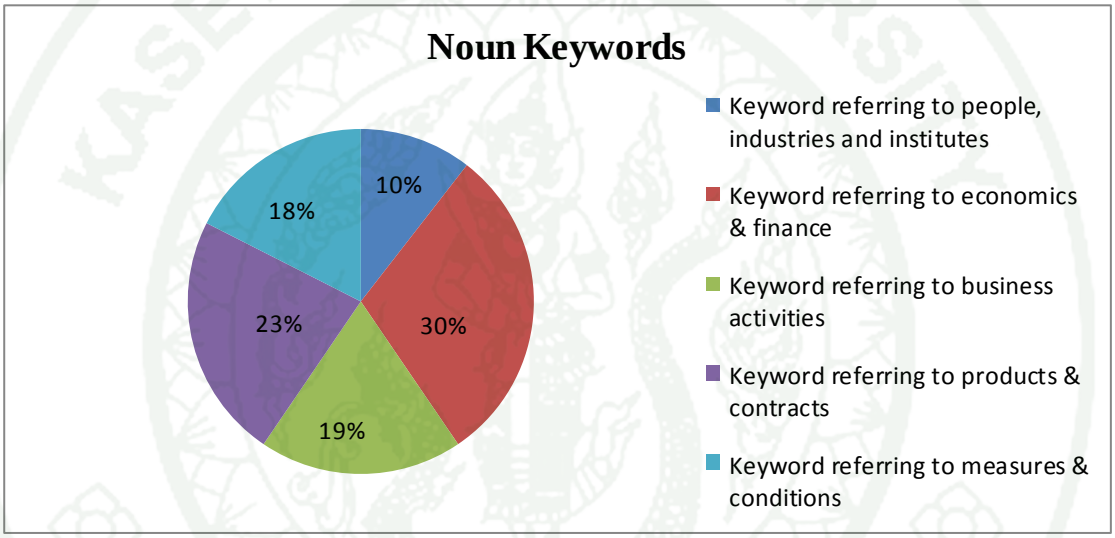

Figure 5.2 Classification of the 58 noun keywords into five areas

The above figure provides a summary of the classification of the total 58 noun keywords. This figure reveals that the majority of the noun keywords are words that refer to economics & finance (30%) and products & contracts (23%), followed by keywords referring to business activities (19%), keywords referring to measures & conditions (18%), and keywords referring to people, industries and institutes (10%). The full list of the classification of the 58 noun keywords is presented in Table 5.2**.**

Table 5.2 Classification of the 58 noun keywords

People, Industries and Institutes	**Economics & Finance**	**Activities**	**Products & Contracts**	**Measures & Conditions**
Insurer	Demand	Insurance	Health	Model
Firm	Value	Risk	Life	Level
Individual	Price	Reinsurance	Coverage	Measure
Company	Variable	Market	Policy	Year
Consumer	Ratio	Analysis	Liability	Period
Reinsurer	Probability	Distribution	Property	Function
	Wealth	Management	Loss	Factor
	Rate	Plan	Premium	Cost
	Asset	Information	Claim	Equilibrium
	Hedge	Selection	Contract	Annuity
	Stock	Rating	Benefit	
	Capital		Term	
	Investment		Case	
	Income		Group	
	Bond			
	Utility			
	Return			

Regarding the prominent role of the 58 noun keywords and the possibility of categorising them into different areas, this allows a better understanding of the general idea of insurance. In this study, it can be seen that the people involved in the insurance field include *insurer*, *consumer*, and *reinsurer*. Products such as *policy*, *coverage*, *property*, and *liability* are essential to their activities. They participate in business activities which are related to *risk* and *insurance*. The activities they engage in are related to *possibility*, *demand*, *price*, and *rate*, with key measures and conditions such as *factor*, *period*, and *year*.

Adjectives

The next word class to be considered is adjectives. Out of 64 keywords, three are adjectives. It is possible to put them into two categories. The first category consists of an item that directly related to products of insurance: *private*, e.g. *private insurance*, *private health insurance*, and *private passenger auto liability*. The second

category emphasises activities involved with finance and money: *financial* and *optimal*. *Financial* displays strong meanings related to finance since the word itself functions as an adjective of the word 'finance'. However, the word *financial* found here conveys genre-specific meanings to insurance as can be seen in phrases such as *financial risk management* and *financial and actuarial*. *Optimal*, likewise, displays a positive semantic sense of activities in insurance related to money and finance. For example, *optimal asset allocation* refers to the allocation of assets in an investment portfolio in which the investors have to consider the possibility of risks. In trading there are risks of price fluctuation, and *optimal hedging* is a method of reducing risk of loss involving contracts of commodities bought and sold at certain periods of time with certain prices.

Verbs

From the list of the top 64 keywords, the number of verbs found is minimised, only one. In a business environment, it is not beyond expectation to find a verb which is associated with business: *pay*. The verb *pay* is used in context and displays meanings related to insurance activity. The strong connection between *'pay'* and its collocates *'claims'* and *'losses'* is noticeable. The collocates *claims*, *losses*, *premium* and *insurance* allow the possibility of drawing a line of semantic association to insurance liability and coverage of the verb *pay*. Illustrated below is a concordance example:

```
1     sufficient liquidity to pay claims at any time, without
2  willingness and ability to pay claims. Furthermore, an insurer
3     to provide coverage and pay claims. Although 2006 was a calm
4 reinsurer is only liable to pay losses over a defined amount
5          Fund obligation to pay losses to the sum of its assets
6      return only willing to pay a premium for the risk that is
7       and low severity risk pay a premium that is higher than their
8 individual airlines have to pay for insurance. Underwriters will
9   individual willingness to pay for insurance and the expected loss
```

The above concordances suggest that ESP students in the field of insurance should learn possible collocations of the verb *'pay'*. In insurance, it is very likely to

find phrases such as *pay claims*, *pay losses*, *pay a premium*, and *pay for insurance.* Rather than studying the word alone, learners should be exposed to these useful phrases as they are leaning English for insurance.

Nouns/ Adjectives

The duality of word class is another aspect to be considered in word assignation. In the present study *average* (56) as a noun displays strong meanings which show the connection to insurance measures and contacts. An example in the concordance below shows how *average* collocates with measure-related clusters.

```
1    below its historical average. In our preliminary results, these
2claims is 8.4 percent on average. The selection model provides a
3    insurance purchases. Average income also increases monotonically
4      decreases with the average life expectancy (lower probability
5    buy insurance at the average premium rate? Begin by noting that
```

From the above concordance, *average* functions as a noun in the cluster *historical average* which in context talks about the average rate of return of the equity investment in the past. As a noun, it is found the *average* is used in the pattern of number + percent + preposition (on) + average, as in *8.4 percent on average*. This pattern is potentially useful to ESP students for creating an accurate pattern of the word *average* that is found with high frequency in the field of insurance.

Average also appears as an adjective. Sample clusters from the IRAC presented in the above concordance show how *average* collocates with money-related terms, as can be seen in *average income* which in context reflects the purchasing power of consumers when their *average income* is higher. *Average* is concerned with price, as in *average premium rate*. In life insurance, the possibility of death creates a great deal of demand for life insurance products. This may be explained by the finding of the cluster *average life expectancy*. These examples and explanation of the usage of the word *average* have hopefully given a clearer picture of its usage as a noun and an adjective in the insurance context.

Nouns/ Verbs

The remaining word is *purchase* (42), which is closely associated with insurance activities. From the IRAC, *purchase* as a verb collocates with the insurance product-related cluster *purchase full coverage*, which refers to the policy that insureds are fully covered against every exposure to loss. *Purchase* as a noun is found in the clusters such as *mandatory purchase*. Learners may need to know that insurance policy is either mandatory or voluntary. Mandatory is one that is enforced by law, such as auto insurance, while health insurance is not. Having seen the use of *purchase* with its collocate *mandatory*, there might be a question if it can be replaced by the word 'obligatory' since they have similar definitional meanings. From the use of concordancing software, it is found that there is no hit of a cluster that uses the word 'obligatory' preceding 'purchase', and that makes it possible to say that the cluster 'obligatory purchase' is unacceptable.

So far, it has been seen that the content keywords discussed above can reflect people, industries and institutes, economy and finance, activities, products and contracts, measures and conditions in the insurance field. In order to get a clearer picture of the keywords of the IRAC, further discussion can also be presented through the distribution of keywords according to what types of vocabulary they belong.

The top 64 keywords are distributed into groups according to their difficulties. The result of this analysis shows that keywords with semantic meaning related to insurance are spread over different types of vocabulary lists. The full details are presented in Table 5.3. Interestingly, keywords like *insurance* which are rated to step 3 as word with specific meaning to insurance are not in the OFF-LIST, which should be strongly related, but instead appear in the GSL. This is because GSL is based merely on frequency of occurrences in general text. This then highlights the significance of the word 'insurance', which belongs to both the basic wordlist that learners need to know and also the key wordlist that they should know. Yet never before has anybody researched the lexical behaviours of word 'insurance' in terms of its collocates and how it is used in context.

Table 5.3 Vocabulary wordlist distribution of the top 64 keywords

Keywords	Total No.	Wordlist		Frequent/ Easy
average, capital, case, claim, company, cost, coverage, demand, group, level, life, loss, market measure, pay, plan, price, private, property, rate, rating, return, stock, term, value, wealth, year firm, health, information, insurance, management, model, probability, risk	**35**	**GSL**	**55%**	
analysis, benefit, bond, consumer, contract, distribution, factor, financial, function, income, individual, investment, period, policy, purchase, ratio, selection, utility, variable	**19**	**AWL**	**30%**	
annuity, asset, equilibrium, hedge, insurer, liability, optimal, premium, reinsurance, reinsurer	**10**	**OFF-LIST**	**15%**	**Rare/ Difficult**

The finding clearly indicates that appearing in the OFF-LIST are words that have technical meanings in the insurance field such as *insurer*, *reinsurance*, and *reinsurer*. It is also found that keywords with meanings related to insurance are not necessarily non-GSL & AWL. Keywords to insurance are also shared and used in other disciplines as it has been found that *"all disciplines adapt words to their own ends, displaying considerable creativity in both shaping words and combining them with others to convey specific, theory-laden meanings associated with disciplinary models and concepts"* Hyland and Tse (2007: 245-246).

Even though there are as many as 35 words out of the 64 keywords (55%) that belong to the General Word List or GSL, this does not mean that these 35 words are easy and simple words because they can be used in a specific meaning. That is why we need to find out how they are used in context and which words collocate with them.

An example mentioned in Table 5.3 is the use of the word *claim*. As a general simple word, it displays two meanings found in COBUILD. One of the meanings with its usage as a noun simply means "*a demand for something that you have a right to"*. The second meaning as a verb is widely used with meaning as defined *"If you claim something, you try to get it because you think you have a right to it".* The meanings of the word *claim* in general senses are not difficult to guess or understand because of its high frequency in usage.

In the IRAC the term *claim* appears very frequently as it has meaning related directly to insurance. It collocates with words reflecting its technical use in the insurance field such as in *claim containment costs*, *claim deductible system*, and *claim settlement costs*. These clusters specify the activities of the insurance claim system. These examples show the difficulty in guessing the meaning when the word *claim* appears in the insurance environment. People who are not in the field of insurance may not know that 'claim deductible system' means a method that as established in contract, a policyholder must be responsible for the deductible amount for every claim made, or that 'claim settlement costs' are the costs that arise with the claim-processing costs of settlement to the policyholder. Only through a corpus-based such as this study can these useful phrases spring upon us.

From this example comparing usage of the word *claim* between insurance English and general English, the findings prove to be crucial that easy words in the GSL list are not always simple. The words seem to be more difficult and genre-specific when they appear in the insurance context.

In summary, given the procedure of identifying a keyword through the use of corpus-based approach, this study discovers that even though the findings found the list of keywords that seems to be key lexical items of insurance English, it is evident that the keywords found in the IRAC do not necessarily express strong semantic technicality to the insurance field. It can be seen that among the top 100 keywords, through the use of a rating scale approach, there are 34 keywords found that have no meanings related to insurance. Therefore, the use of human intuition as in a rating

scale is remarkably important in judging the meanings of words. Having experience in the field, an expert is able to tell if the words have meanings related to insurance. However, bias could be a major problem of this method. The findings at this point allow the possibility to notice that it could be more effective in finding keywords with specialised meaning to insurance in the IRAC through the application of not only one particular method but instead using both statistical approach as for corpus-based approach and intuition of experts in the field of insurance as for rating-scale approach.

Most Frequent Collocations of the Keywords Found in the IRAC

The present study has provided a list of the twenty five most frequent collocates of 100 keywords presented in the manner of 2-word clusters, 3-word clusters and 4-word clusters. Among those lists, it is found that a smaller cluster is often a part of a bigger cluster (Sinclair, 1991). The present study has found a similar pattern as shown in Figure 5.3.

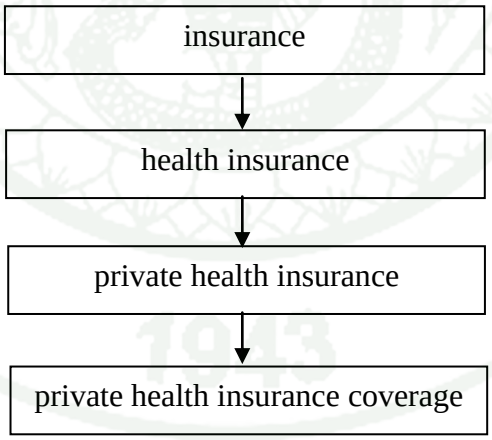

Figure 5.3 Hierarchy of lexical bundles with ‘insurance’ as the node

In Figure 5.3, the node *insurance* occurs first with the collocate *health* to form the collocation of the 2-word cluster *health insurance*. This cluster appears 651 times. The number of occurrences is lower when the combination is enlarged to the 3-word

cluster *private health insurance*, which occurs 82 times, and the 4-word cluster *private health insurance coverage*, which appears 13 times. From this example, we see a clear semantic association between *"private"*, *"health"*, *"insurance"* and *"coverage"*. The connection requires knowledge in the insurance area to understand that this is one of many types of insurance products. In a specialised area, abbreviation is also seen as a key element. In the present study, the node *bond* collocates with the abbreviation *CAT* as in *CAT bond*. To understand the meaning of this collocation, specific understanding of the field is required. It is significant for learners to know that *CAT* stands for "Catastrophe", so CAT bond actually refers to a kind of bond that insurance companies issue to investors in order to raise funds in case a catastrophe occurs.

Regarding the frequency of occurrences, and after applying the criteria of the cutoff point, the finding reveals the larger number of 2-word clusters than 3- and 4-word clusters in the IRAC. This finding supports Shin and Nation's (2008) claim that the shorter the collocation, the greater the frequency.

Pedagogical Implication

Creating a list of frequently used collocations of keywords in the field of insurance is one of the main purposes of this present study. It is to aid ESP learners and people who work in the field to understand better the collocational items related to the field they are involved in. It is undeniable that vocabulary is an important need for employees in the insurance industry (Sangvaree, 2004), but the study of it is limited. The list of keywords and their collocates as shown in Appendix can be used as a reference for English teachers who can now teach collocation in English in the field of insurance. It has been remarked that it is vitally important for learners in a specific field to not only know the jargon or technical terminology which they know how to use, but also to comprehend the collocations of sub-technical vocabulary (Lewis, 2000).

Therefore, it is believed that, with respect to increasing lexical proficiency and expanding and developing vocabulary repertoire for learners, the list of collocations of the top 64 keywords found in the IRAC will be useful to facilitate vocabulary acquisition.

More importantly to the teaching aspect, the result of collocational clusters found in this study can be adapted and used as classroom material, as the results are obtained from an authentic source where the language used in research articles is highly accepted and recommended for academic purposes. To ensure that students are well-exposed to the target collocation feature, teachers can show authentic examples to students by using a concordance tool so that students can see the context and its usage in the insurance context.

To help English students acquire vocabulary effectively, the Lexical Approach (Lewis, 1997) has raised awareness of how important it is for learners to recognise and be able to use lexical structures in the forms of clusters or chunks of words. Having recognised the keywords in insurance as well as their collocational behaviours, the findings of this research allow for practical application in creating a collection of exercises focused on the use of collocation in insurance field. Using authentic resources in the collocation exercises from target corpora is not exactly a new idea; prominent examples can be seen in such works as *English Collocations in Use (Advanced)* (O' Dell and McCarthy, 2008) and *Collocations Extra* (Walter and Woodford, 2010). However, these books focus on the general use of collocations instead of in any particular fields. With respect to the need to learn vocabulary in the insurance field as suggested by Sangvaree (2004), this study has created a set of collocation exercises based on the presented outcome of the concordances shown by the Concordance tool *AntConc* (Anthony, 2007). Only the top three keywords ‘insurance’, ‘risk’, ‘insurer’ and their collocates are selected to be included in the four different types of exercises: (1) Collocations Match-UP, (2) 3 – 4 Clusters Challenge, (3) Collocations Hunting, and (4) Skill Practice.

1. Collocations Match-UP: Match the two parts of these collocations

1. primary	insurance
2. risk	aversion
3. health	insurer

2. 3 – 4 Clusters Challenge: Complete the collocations using the words below

insurance risk insurer

1. private health ……
2……...management and insurance
3……and reinsurer

3. Collocations Hunting: Choose the correct collocation

1 Over the course of multiple rounds, the potential insureds attempt to purchase …… from the potential insurers.

a. *insurance* b. cost c. loss d. premium

2. Companies with very low risk exposures solve their……..... management issues without external help and rely on brokers only for transaction services.

a. investment b. internal c. damage d. *risk*

3. Stockholders supply capital to a stock ……only when they expect to earn a fair return on their investment.

a. money b. firm c. *insurer* d. model

4. Skill Practice: Complete each sentence using the given words below

insurance risk insurer

1. Some people may lack private health……coverage during a particular time period because they become frictionally uninsured. (insurance)

2. ……..**aversion** refers to the tendency to prefer to pay a defined sum of money that is known with certainty instead of being exposed to the risk of suffering a larger and uncertain financial loss in the future. (risk)

3. To back up their promises to pay claims under the terms of their policies, both the **primary**……and the reinsurer hold equity capital to provide sufficient funds in case of adverse loss or investment fluctuations. (insurer)

The above set of collocation exercises are ordered in degree of difficulty. Learners could start off easy by matching 2-word clusters (Exercise 1) to raise awareness of collocational patterns before moving on to a longer set of 3- and 4- word clusters, respectively (Exercise 2). Exercises 3 and 4 are more difficult and challenging for learners, who will need to apply their cognitive skill in reading and inferring meanings from context. Vocabulary exercises are most fruitful for learners only when they are designed to be recyclable and ordered in steps of difficulty which can help scaffold learners during their learning stages. It is hoped that the suggested exercises may provide some ideas and guidelines for teachers to use or adapt in the classroom to exercise the learners' repertoire of vocabulary in the insurance field.

Limitations and Recommendations for Further Study

The findings of this study also reveal the vital role of implementing a corpus-based approach in vocabulary learning in which learners can gain access to authentic material and actual usage in context. Teaching concordances is yet another issue that is worth researching, but this is beyond the scope of this study.

The BNC Sampler was the only reference corpus used in comparison to the target corpus IRAC here. Therefore, future researchers could consider using a larger reference corpus or selecting more corpora which may be more effective in terms of finding frequency and keyness. While the present study has put more emphasis on providing the list of frequently used keywords and their collocations, other aspects of language features and usage have not fallen within the scope of this study. It is hoped,

however, that the lexical findings in this study will prove valuable for ESP learners who study in the insurance field.

Summary of the Thesis

This present study was carried out with an aim to identify and analyse a list of most frequent keywords and their collocations. To achieve this, a corpus-based approach was employed. The IRAC was created to compare with the reference corpus, the 1 million written words of the BNC Sampler corpus. The electronic versions of published academic articles were collected, compiled and analysed by means of the software tool *AntConc* with the log-likelihood statistical test of lexical keyness used to detect keywords and their collocational patterns.

The list of the top 64 keywords and their collocations certainly has fundamental importance to the study of English in the field of insurance. It is worth introducing these frequently used lexical items to ESP students in the insurance field and others who are involved in this industry to learn, and for teachers to teach. It is recommended that the keywords and their collocational behaviours found in this study as shown in Appendix be incorporated into the vocabulary lessons of insurance as well as teaching materials. Through suggested vocabulary exercises discussed earlier in this section, learners can become more familiar with the lexical collocations of important keywords frequently used in the insurance field.

REFERECES

Anthony, L. 2004. "AntConc: A learner and classroom friendly, multi-platform corpus analysis toolkit." In **Proceedings of IWLeL 2004. An interactive workshop on Language e-Learning.** Tokyo: Waseda University, 7-13.

Anthony, L. 2007. **AntConc 3.2.1w (Windows)** (Online). www.antlab.sci, December 25, 2011.

Bailey, S. 2006. **Academic Writing: A Handbook for International Students**. London: Routledge.

Baker, P. 2004. "Quering keywords: Questions of Difference, Frequency, and Sense in Keywords Analysis." **Journal of English Linguistics** 32 (4): 346-359.

Biber, D. 1993. "Representativeness in corpus design." **Literary and linguistic computer** 8 (4): 243-257.

_____. 1995. **Dimensions of Register Variation**. Cambridge: Cambridge University Press.

Bowker, L. and J. Pearson. 2002. **Working with specialized language: a practical guide to using corpora**. London: Routledge.

Brett, P. 1994. "A genre analysis of the results section of sociology articles." **English for Specific Purposes** 13 (1): 45-59.

Bruce, I. 2008. "Cognitive genre structures in Methods sections of research articles: A corpus study." **Journal of English for Academic Purposes** 7 (1): 38-54.

_____. 2009. "Results sections in sociology and organic chemistry articles: A genre analysis." **English for Specific Purposes** 28 (2): 105-124.

Chalinrat, J. 2001. **Readability of English Non-Life Insurance Policies**. Master of Arts in Communicative English, Ramkhamheng University.

Chen, Q. and G. Ge. 2007. "A corpus-based lexical study on frequency and distribution of Coxhead's AWL word families in medical research articles (RAs)." **English for Specific Purposes** 26: 502–514.

Chung, T. M. and P. Nation. 2004. "Identifying technical vocabulary." **System** 32: 251- 263

Cohen, L., L. Manion, and K. Morrison. 2000. **Research Methods in Education.** 5th ed. New York: Routledge Falmer.

Cortes, V. 2004. "Lexical bundles in published and student disciplinary writing: Examples from history and biology." **English for specific purposes** 23: 397-423.

Coxhead, A. 2000. "A new academic wordlist." **TESOL Quarterly** 34 (2): 213–238.

Crystal, D. 2003. **A Dictionary of Linguistics and Phonetics**. Oxford: Blackwell Publishers.

Dunning, T. 1993. "Accurate Methods for the Statistics of Surprise and Coincidence." **Computational Linguistics** 19 (1): 61-74.

Durrant, P. 2008. **High-frequency collocations and second language learning** (Online). http://etheses.nottingham.ac.uk/622/, December 30, 2011.

_____. 2009. "Investigating the viability of a collocation list for students of English for academic purposes." **English for Specific Purposes** 28: 157-169.

Firth, J. R. 1957. "Modes of Meaning." In J. R. Firth. (ed.). **Papers in linguistics 1934-1951.** Oxford: Oxford University Press, 190-215.

Flowerdew, L. 1998. "Corpus Linguistic Techniques Applied to Textlinguistics." **System** 26: 541-552.

Flowerdew, L. and M. Peacock. (eds.). 2001. **Research perspectives on English for academic purposes.** Cambridge: Cambridge University Press.

Francis, G. and J. Sinclair. 1994. "'I bet he drinks Carling Black Label': a riposte to Owen on corpus grammar." **Applied Linguistics** 15 (2): 190-200.

Gledhill, C. 2000a. **Collocations in science writing.** Tübingen: Gunter Narr Verlag.

_____. 2000b. "The discourse function of collocation in research article introductions." **English for Specific Purposes Volume** 19 (2): 115-135.

Halliday, M. A. K. 1966. "Lexis as a linguistic level." In C. E. Bazell, J. C. Catford, M. A. K. Halliday, and R. H. Robins. (eds.). **In Memory of J. R. Firth**. London: Longmans, Green and Co. Ltd., 148-162.

Harter, P. S. 1998. "Scholarly Communication and Electronic Journals: An Impact Study." **Journal of the American society for information science** 46 (6): 507-516.

Hathaway, G. H. and P. A. Marroso. 2001. **Plain English in Insurance Paper** (Online). www.michbar.org/generalinfo/plainenglish/columns/94_feb.html, January 5. 2011.

Hill, J. 2000. "Revising priorities: From grammatical failure to collocational success." In M. Lewis. (ed.). **Teaching collocation: Further developments in the lexical approach**. Hove, England: Language Teaching Publications, 47-69.

Hoey, M. 1991. **Patterns of Lexis in Text**. Oxford: Oxford University Press.

Holbrook, E. 2012. The 10 Largest Risk Management and Insurance Education Programs. In **Risk Management** (Online). http://rmmagazine.com, September 22, 2012.

Holmes, R. 1997. "Genre Analysis and the Social Sciences: An investigation of the Structure of Research Article Discussion Sections in Three Disciplines." **English for Specific Purposes** 16 (4): 321-337.

Hopkins, A. and T. Dudley-Evans. 1998. "A genre-based investigation of the discussion sections in articles and dissertations." **English for Specific Purposes** 7 (2): 113-121.

Hyland, K. 2008. "Academic clusters: Text patterning in published and postgraduate writing." **International Journal of Applied Linguistics** 18 (1): 41-62.

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

Jones, S. and J. M. Sinclair. 1974. "English lexical collocations. A study in computational linguistics." **Cahiers de Lexicologie** 24: 15–61.

Kennedy, G. 1998. **An Introduction to Corpus Linguistics**. Harlow: Addison Wesley Longman.

Klein-Wohl, E. and A. Yahav. 2000. **Reading English with awareness for academic studies**. 2 vols. Tel-Aviv: The Open University of Israel.

Kushner, S. 2003. "Tackling the Needs of Foreign Academic Writers: A Case Study." In Beer, D. F. (ed.). 2nd ed. **Writing and speaking in the technology professions: a practical guide**. New Jersey: John Wiley & Sons, Inc.

Leech, G. 1991. "The State of the Art in Corpus Linguistics." In Aijmer, K. and Altenberg, B. (eds). **English Corpus Linguistics, Studies in Honour of Jan Svartvik**. London: Longman.

Lewis, M. 1997. **Implementing the Lexical Approach: Putting Theory into Practice.** Hove: Language Teaching Publications.

Lewis, M. (ed.). 2000. **Teaching collocation: Further development in the lexical approach**. Oxford: Oxford University Press.

Lia, Y. and D. D. Qian. 2010. "Profiling the Academic Word List (AWL) in a financial corpus." **System** 38 (3): 402-411.

Makins, M. (ed.). 1991. **Collins English Dictionary.** 3rd ed. Glasgow, Harper Collins.

Marco, M. J. L. 2000. "Collocational frameworks in medical research papers: A genre- based study." **English for Specific Purposes** 19: 63–86.

Martínez, A. L., C. S. Beck, and B. C. Panza. 2009. "Academic vocabulary in agriculture research articles: A corpus-based study." **English for Specific Purposes** 28: 183–198.

McCarthy, M. 1990. **Vocabulary**. Oxford: Oxford University Press.

Nation, I.S.P. 2001. **Learning Vocabulary in Another Language**. Cambridge: Cambridge University Press.

Nelson, M. A. 2000. **Corpus-Based study of business English and business English teaching materials.** (Electronic version). Unpublished PhD Thesis. Manchester: University of Manchester.

O' Dell, F. and M. McCarthy. 2008. **English Collocations in Use (Advanced)**. Cambridge: Cambridge University Press.

Peacock, M. 2002. Communicative moves in the discussion section of research articles. **System** 30 (4): 479-497.

Prudential Insurance Company of America, (ed.). 1915. **The Documentary History of Insurance, 1000 B. C. - 1875 A. D.** Newark, N.J: Prudencial Press.

Rizvi, M. Ashraf. 2005. **Effective Technical Communication**. Tata McGraw Hill: New Delhi.

Samraj, B. 2002. "Introductions in research articles: variations across disciplines." **English for Specific Purposes** 21 (1): 1-17.

Sánchez Hernández, P. 2005. "Examining English for academic purposes students' vocabulary output: Corpus-aided analysis and learner corpora." **Volumen Monográfico**. 201-212.

Sangvaree, K. 2004. **A Study of the English Language Used in Insurance Business.** Master of Arts in English for Business and Technology, The University of the Thai Chamber of Commerce.

Scott, M. 1997. "PC analysis of key words – and key key words." **System** 25 (2): 233–245.

_____. 2001. "Comparing corpora and identifying key words, collocations, and frequency distributions through the WordSmith Tools suite of computer programs." In M. Ghadessy, A. Henry, and R. L. Roseberry. (eds.). **Small corpus studies and ELT theory and practice**. Amsterdam: John Benjamins Publishing Company, 47-70.

_____. 2009. **WordSmith tools version 5.0.** Oxford: Oxford University Press.

Shin, D. and P. Nation. 2008. "Beyond single words: the most frequency collocations in spoken English." **ELT Journal** 62 (4): 339-348.

Sinclair, J. M. 1966. "Beginning the study of lexis." In C. E. Bazell, J. C. Catford, M. A. K. Halliday, and R. H. Robins. (eds.). **In memory of J.R. Firth.** London: Longman, 410-430.

_____. 1991. **Corpus, Concordance, Collocation**. Oxford: Oxford University Press.

_____. (ed.). 2001. **Collins COBUILD English Dictionary for Advanced Learners.** 3rd ed. London: HarperCollins Publishers.

Sinclair, J. M. (ed.). 2003. **Collins COBUILD Advance Learner's English Dictionary.** 4th ed. Glasgow: Harper Collins Publishers.

_____., S. Jones, and R. Daley. 2004. **English Collocation Studies: The OSTI Report**. London & New York: Continuum.

Sudsa-Ard, A. 2001. **An Analysis of English Linguistic Forms in a Marine Cargo Insurance Policy**. Master of Arts Thesis in Communicative English, Ramkhamhaeng University.

Swales, J. 1990. **Genre analysis: English in academic and research settings**. Cambridge: Cambridge University Press.

Thurstun, J. and C. N. Candlin. 1998. "Concordancing and the teaching of the vocabulary of Academic English." **English for Specific Purpose** 17 (3): 267–280.

Vongpumivitch, V., J. Y. Huang, and Y. C. Chang. 2009. "Frequency analysis of the words in the Academic Word List (AWL) and non-AWL content words in applied linguistics research papers." **English for Specific Purposes** 28: 33-41.

Walter, E. and K. Woodford. 2010. **Collocations Extra. Multi-level Activities for Natural English**. Cambridge: Cambridge University Press.

Ward, J. 2007. "Collocation and technicality in EAP engineering." **Journal of English for Academic Purposes** 6: 18–35.

_____. 2009. "A basic engineering English word list for less proficient foundation engineering undergraduates." **English for Specific Purposes Journal** 28 (3): 170-182.

West, M. 1953. **A general service list of English words**. London: Longman.

APPENDIX

1. Comprehensive list of the 2-, 3- and 4-word clusters keyword collocations

No.	Word	2-Word Clusters		3-Word Clusters			4-Word Clusters		
		on the right	on the left						
1	insurance	insurance market	life insurance	management and	insurance		risk management and	insurance	
		insurance company	health insurance	demand for	insurance		demand for life	insurance	
		insurance coverage	purchase insurance	risk and	insurance		[3]association of	insurance	commissioners
		insurance industry	liability insurance	private health	insurance		life	insurance	and annuities
		insurance demand	private insurance	health	insurance	coverage	demand for health	insurance	
		insurance premium	auto insurance		insurance	and reinsurance	long term care	insurance	
		insurance contract	buy insurance	life	insurance	demand	management and	insurance	research
		insurance policy	automobile insurance	property casualty	insurance		national association of	insurance	
		insurance business	flood insurance	life	insurance	market	amount of life	insurance	
		insurance purchase	homeowners insurance	property liability	insurance		devoted to health	insurance	
		insurance journal	casualty insurance	level of	insurance		life	insurance	and annuity
		insurance regulation	nonlife insurance	dwell fire	insurance		private health	insurance	coverage
		insurance research	property insurance	life	insurance	products	dwelling fire	insurance	policies
		insurance sector	medical insurance	life	insurance	purchases		insurance	and reinsurance markets
		insurance student	fire insurance	[1]medical malpractice	insurance		optimal level of	insurance	
		insurance line	full insurance	purchase of	insurance		property and casualty	insurance	
		insurance plan	coverage insurance	cost of	insurance				
		insurance article	competitive insurance	[2]lines of	insurance				
		insurance literature	malpractice insurance	association of	insurance				
		insurance agent	terrorism insurance	life	insurance	industry			
		insurance program	participating insurance	types of	insurance				
		insurance product	buying insurance	health	insurance	articles			
		insurance consumption	group insurance	amount of	insurance				
		insurance buyer	public insurance	full coverage	insurance				
		insurance pricing			insurance	and annuities			

[1] *medical malpractice* insurance is one of the insurance products which provides coverage for professional liability (e.g. physicians and surgeons).
[2] in insurance business there are four major *lines of insurance*, including property, casualty, life, and health and disability.
[3] the cluster association of *insurance commissioners* is used here to refer to, 'The National Association of Insurance Commissioners', the U.S. standard-setting and regulatory support organization.

2	risk	risk management	low risk		risk	and insurance		risk	management and insurance
		risk aversion	high risk	[1]aggregate mortality	risk		coefficient of	risk	aversion
		risk free	mortality risk		[2]risk	minimization hedging		risk	minimization hedging using
		risk type	asset risk	[3] relative	risk	aversion	constant relative	risk	aversion
		risk minimization	background risk		[4]risk	free rate	high general	risk	policyholders
		risk measure	relative risk	coverage	risk	correlation		risk	of fair contracts
		risk classification	higher risk	low	risk	type	low general	risk	policyholders
		risk based	coverage risk	value at	risk		general	risk	policyholders purchase

No.	Word	2-Word Clusters		3-Word Clusters			4-Word Clusters		
		on the right	on the left						
		risk neutral	lower risk		risk	based capital	high and low	risk	
		risk exposure	severity risk	absolute	risk	aversion	levels of	risk	aversion
		risk premium	credit risk	high	risk	areas	high	risk	indifference curve
		risk adjusted	systemic risk	low	risk	areas	low	risk	indifference curve
		risk policyholders	frequency risk	general	risk	policyholders	degree of	risk	aversion
		risk group	absolute risk	level of	risk		level of	risk	aversion
		risk taking	catastrophic risk	low	risk	individuals		risk	and insurance association
		risk correlation	basis risk		risk	and return			
		risk customers	general risk	coefficient of	risk				
		risk area	different risk	high	risk	individuals			
		risk reduction	longevity risk	interest rate	risk				
		risk tolerance	catastrophe risk	corporate	risk	management			
		risk class	less risk		risk	management action			
		risk estimates	shortfall risk						
		risk adjusted	underwriting risk						
		risk sharing	market risk						
	[1] *aggregate mortality risk* is used to refer to the prediction of the mortality rate.								
	[2] *risk minimization hedging* refers to a measure used in financial trading strategy.								
	[3] *relative risk aversion* is a measure that concerns the uncertainty of gaining vs. losing investment								
	[4] *risk free rate* is a strategy in risk management which concerns the risk of minimum return on investment.								
3	insurer	insurer financial	life insurer	international	insurer	participation	life	insurer	stock returns
		insurer participation	stock insurer		insurer	stock returns		insurer	sided adverse selection
		insurer performance	international insurer		insurer	investment policy	volatility of an	insurer	
		insurer stock	liability insurer	life	insurer	stock	firm and the	insurer	
		insurer size	primary insurer	[1]	insurer	solvency situation	number of	insurer	insolvencies
		insurer solvency	health insurer	property liability	insurer		sensitivity of life	insurer	
		insurer sided	change insurer		insurer	profit margins			
		insurer default	multiline insurer	life	insurer	equity			
		insurer equity	new insurer		insurer	and reinsurer			
		insurer capital	foreign insurer						
		insurer data	neutral insurer						
		insurer failures	single insurer						
		insurer makes	study insurer						
		insurer offers	initial insurer						
		insurer profit	particular insurer						
		insurer profitability	private insurer						
		insurer believes	deviating insurer						
		insurer investment	incumbent insurer						
		insurer losses	individual insurer						
		insurer pays	mutual insurer						
		insurer risk	sponsoring insurer						

No.	Word	2-Word Clusters		3-Word Clusters			4-Word Clusters		
		on the right	on the left						
		insurer defaults							
		insurer groups							
		insurer insolvency							
	[1] *insurer solvency situation*: a situation that an insurer needs to meet financial obligation to cover liability.								
4	market	market share	insurance market	life insurance	market		medical malpractice insurance	market	
		market value	capital market	cat bond	market		equilibrium in a	market	
		market concentration	stock market	stock	market	response		market	for medical malpractice
		market based	reinsurance market	auto insurance	market			market	value of equity
		market conditions	financial market	capital	market	line	percent of the	market	
		market response	individual market	health insurance	market			market	concentration and profitability
		market risk	group market	private debt	market			market	based discount rate
		market discipline	private market	individual	market	sample	competitive health insurance	market	
		market participants	secondary market	financial	market	returns			
		market price	bond market	automobile insurance	market				
		market structure	residual market	foreign	market	share			
		market sample	foreign market	level of	market				
		market efficiency	competitive market	competitive insurance	market				
		market line	labor market	malpractice insurance	market				
		market returns	bull market						
		market characteristics	debt market						
		market profitability	current market						
		market crash	national market						
		market microstructure	host market						
		market model	same market						
		market shares	product market						
		market valuation	voluntary market						
		market reaction	equity market						
		market oriented	annuity market						
		market competition	domestic market						
5	model	model framework	feedback model		model	of insurance	capital asset pricing	model	
		model parameters	basic model	asset pricing	model			model	of insurance demand
		model includes	pricing model	adverse selection	model			model	of employer behaviour
		model risk	selection model	jump diffusion	model			model	for the insurance
		model using	simple model		model	of participating	results of the	model	
		model predicts	empirical model	develop a	model		two state insurance	model	
		model results	regression model	inheritance decision	model				
		model based	standard model	fixed effects	model				
		model developed	period model						
		model estimation	original model						
		model presented	theoretical model						

No.	Word	2-Word Clusters		3-Word Clusters			4-Word Clusters		
		on the right	on the left						
		model specification	effects model						
		model structure	market model						
		model assessment	insurance model						
			statistical model						
			diffusion model						
			decision model						
			economic model						
			equation model						
			extended model						
			internal model						
			risk model						
			price model						
			proposed model						
6	loss	loss coverage	expected loss	probability of	loss		size of the	loss	
		loss ratio	earnings loss	excess of	loss		changes in the	loss	
		loss distribution	aggregate loss	industry	loss	warranties	increase in	loss	coverage
		loss experience	developed loss		loss	reserve estimation	improvement in	loss	coverage
		loss probability	potential loss	changes in	loss		premium to expected	loss	
		loss reserve	financial loss	increase in	loss		probability of a	loss	
		loss development	industry loss		loss	adjustment expenses	case of a	loss	
		loss state	insurance loss		loss	of generality	difference in	loss	probabilities
		loss distributions	catastrophic loss	risk of	loss		value of aggregate	loss	
		loss severity	large loss	developed	loss	ratios	value of the	loss	
		loss frequency	insurable loss				wealth in the	loss	
		loss probabilities	reported loss				year developed	loss	ratios
		loss adjustment	retained loss						
		loss estimates	average loss						
		loss function	higher loss						
		loss shock	insured loss						
		loss warranties	noninsurable loss						
		loss amounts	total loss						
		loss data	wage loss						
		loss amount	damage loss						
		loss indices	different loss						
		loss premium	economic loss						
		loss exposure	insurers loss						
		loss measure	liability loss						
		loss control	welfare loss						
7	firm	firm size	insurance firm		firm	specific factors	value of the	firm	
		firm specific	new firm		firm	specific characteristics	take over the	firm	
		firm characteristics	individual firm	relation between	firm		size of the	firm	

No.	Word	2-Word Clusters		3-Word Clusters			4-Word Clusters		
		on the right	on the left						
		firm performance	financial firm		firm	fixed effects		firm	and the insurer
		firm level	breached firm	associated with	firm		worker	firm	size exclusion
		firm value	same firm	control for	firm				
		firm age	large firm		firm	level data			
		firm risk	specific firm	measure of	firm				
		firm type	case firm		firm	size exclusion			
		firm fixed	measure firm		firm	level characteristics			
		firm rating	old firm	measure	firm	size			
		firm leverage	post firm	worker	firm	size			
		firm ownership	worker firm						
		firm properties							
		firm specific							
		firm assets							
		firm cash							
		firm data							
		firm earnings							
		firm effects							
		firm financial							
		firm managers							
		firm owners							
8	premium	premium growth	insurance premium	stock risk	premium			premium	to expected loss
		premium income	pure premium	pure	premium	rate	effect of	premium	growth
		premium rate	risk premium	actuarially fair	premium		equity	premium	of stocks
		premium subsidies	fair premium	fair	premium	rate	percent of the	premium	
		premium paid	average premium	stop loss	premium				
		premium revenue	[1] annual premium	employee	premium	contribution			
		premium payment	[2] single premium	net	premium	written			
		premium written	total premium		premium	per ship			
		premium increases	employee premium		premium	indemnity space			
		premium line	net premium	total	premium	income			
		premium indemnity	pooled premium	pay a	premium				
		premium volume	return premium		premium	per unit			
		premium contribution	equity premium	annual	premium	payment			
		premium model	fair premium	average	premium	rate			
		premium savings	written premium	average	premium	savings			
		premium case	additional premium	charge a	premium				
		premium level	loss premium	effect of	premium				
		premium loading	high premium	equilibrium pooled	premium				
		premium taxes	low premium	fair return	premium				
		premium adjustment	market premium	insurance risk	premium				
		premium charged	maximum premium	life insurance	premium				

No.	Word	2-Word Clusters		3-Word Clusters			4-Word Clusters		
		on the right	on the left						
		premium decreases	reinsurance premium	net	premium	revenue			
		premium equals	adjusted premium		premium	payment method			
		premium plus	equilibrium premium						
		premium policy	gross premium						
		premium principle							
	[1] *annual premium*: the amount of money a policyholder pays to an insurance company as a fee to have the insurance coverage.								
	[2] *single premium*: a type of premium payment used in life insurance. An insured entitled to pay one single premium payment for a lifetime of benefits.								
9	cost	cost sharing	total cost		cost	of capital		cost	of public funds
		cost function	high cost		cost	of insurance	amount of	cost	sharing
		cost efficiency	low cost		cost	of equity		cost	of equity capital
		cost growth	hedging cost		cost	per ship		cost	of risk management
		cost containment	average cost		cost	of public		cost	and revenue efficiency
		cost reduction	replacement cost		cost	of risk		cost	of health care
		cost allocation	transaction cost	amount of	cost		social	cost	of public
		cost structure	marginal cost	high	cost	patients		cost	of the guarantee
		cost penalty	audit cost		cost	sharing parameter	defense and	cost	containment
		cost savings	operating cost	sunk	cost	penalty	increase in the	cost	
		cost perspective	prevention cost		cost	of health	percent of the	cost	
		cost patient	fixed cost		cost	and revenue			
		cost policyholders	insurance cost		cost	of providing			
		cost threshold	sunk cost		cost	sharing elasticity			
		cost associated	capital cost		cost	sharing rule			
		cost control	common cost	cover the	cost				
		cost economics	full cost	annual yield	cost				
		cost reduction	initial cost	bear the	cost				
		cost adjustment	overall cost		cost	of auto			
		cost insurance	yield cost		cost	of coverage			
		cost reimbursement	social cost		cost	of medical			
		cost based	agency cost	expected total	cost				
		cost inefficient	opportunity cost						
10	asset	asset allocation	risky asset	optimal	asset	allocation	capital	asset	pricing model
		asset risk	optimal asset		asset	allocation problem		asset	allocation and liquidation
		asset reconstitution	illiquid asset		asset	risk taking	solutions to the	asset	
		asset class	initial asset	risky	asset	reconstitution	optimal	asset	allocation given
		asset return	liquid asset	capital	asset	pricing	case of an	asset	
		asset liability	underlying asset		asset	liability management	change in	asset	demand
		asset price	capital asset		asset	pricing model	insurers	asset	risk taking
		asset demand	different asset	initial	asset	allocation	intangible assets and	asset	
		asset portfolio	multiperiod asset	risk free	asset		using the underlying	asset	
		asset meltdown	riskless asset		asset	allocation decisions	aggregate impact of	asset	

No.	Word	2-Word Clusters		3-Word Clusters			4-Word Clusters		
		on the right	on the left						
		asset composition	free asset		asset	risk factor	annuitization and	asset	allocation
		asset volatility	each asset	riskless	asset	reconstitution		asset	price feedback formulas
		asset value	insurers asset		asset	and liability	impact of	asset	demand
		asset side	financial asset	multiperiod	asset	allocation	related to	asset	risk
		asset accumulation	other asset		asset	allocation given	wealth and	asset	price
		asset correlation	total asset		asset	pricing models			
		asset holding	all asset		asset	reconstitution states			
		asset management	basic asset	different	asset	classes			
		asset base	best asset	insurers	asset	risk			
		asset category	intangible asset		asset	risk measures			
		asset mix	risk asset	capital to	asset				
		asset model	same asset	change in	asset				
			suboptimal asset	related to	asset				
			bank asset		asset	liability matching			
			based asset						
11	rate	rate sensitivity	interest rate		rate	of return	free	rate	of return
		rate risk	growth rate	risk free	rate			rate	of time preference
		rate assumptions	free rate	guaranteed interest	rate		expected	rate	of return
		rate regulation	discount rate	interest	rate	sensitivity	risk free interest	rate	
		rate suppression	marginal rate	interest	rate	risk	marginal	rate	of substitution
		rate volatility	tax rate	discount	rate	assumptions	unexpected interest	rate	changes
		rate compression	unemployment rate		rate	of interest	free	rate	of interest
		rate decreases	inflation rate	tax deduction	rate		based discount	rate	assumptions
		rate environments	premium rate	interest	rate	changes	growth	rate	in net
		rate level	mortality rate	pure premium	rate		growth	rate	of insurance
		rate making	error rate	free interest	rate		market based discount	rate	
		rate regulatory	surrender rate	unexpected interest	rate			rate	in net premiums
		rate changes	deduction rate	marginal tax	rate			rate	of direct losses
		rate increases	expected rate	market interest	rate			rate	sensitivity of life
		rate structure	insurance rate		rate	of substitution	real	rate	of return
		rate variable	wage rate	hourly wage	rate		fair	rate	of return
			death rate	insurance coverage	rate		growth	rate	of premiums
			participation rate	interest	rate	volatility	health insurance coverage	rate	
			real rate		rate	of growth	pension discount	rate	assumptions
			audit rate	annual growth	rate			rate	of return targeting
			coinsurance rate	based discount	rate			rate	on line index
			replacement rate	interest	rate	environments	real	rate	of interest
			uninsured rate	median growth	rate				
			coverage rate		rate	of insurance			
			poverty rate						

No.	Word	2-Word Clusters		3-Word Clusters			4-Word Clusters		
		on the right	on the left						
12	price	price elasticity	gross price		price	of medical		price	of medical care
		price index	house price	house	price	index		price	elasticity of demand
		price increases	stock price		price	of insurance	gross	price	of medical
		price returns	fair price	gross	price	elasticity	increase in the	price	
		price regulation	market price	house	price	returns	[2] nationwide house	price	index
		price change	fixed price	fixed	price	contract		price	elasticity of medical
		price dynamics	asset price	real world	price		real world	price	dynamics
		price line	cash price	stock	price	returns	change in the	price	
		price volatility	transfer price	fair	price	line	medical care gross	price	
		price contract	underlying price	[1] insurance	price	regulation	asset	price	feedback formulas
		price model	sales price	nationwide house	price		care gross	price	elasticity
		price process	insurance price		price	and income		price	of health insurance
		price discount	net price		price	and quantity	wealth and asset	price	
		price equal	futures price		price	of health			
		price reaction	option price		price	of reinsurance			
		price series	world price	world	price	dynamics			
		price curve	consumer price	consumer	price	index			
		price data	care price	asset	price	processes			
		price differentials	mid price	care gross	price				
		price evolution	reservation price	changes in	price				
		price inflation	average price	fair	price	curve			
		price level	competitive price	medical care	price				
		price quotes	equilibrium price	stock	price	reaction			
		price feedback	initial price	asset	price	feedback			
			maximum price	current market	price				

[1] *insurance price regulation*: the price of insurance market in the U.S. that is regulated by the state regulatory authorities through the coordination with the National Association of Insurance Commissioners (NAIC).

[2] *nationwide house price index*: information on house prices nationwide. The value of the house has a significant impact on insurance premiums.

No.	Word	2-Word Clusters		3-Word Clusters			4-Word Clusters		
		on the right	on the left						
13	value	value function	present value		value	at risk		value	of the company
		value approach	market value	net present	value			value	of the firm
		value associated	expected value	insurance surrender	value		market	value	of equity
		value care	extreme value	extreme	value	distributions	reductions in the	value	
		value distribution	fair value		value	of equity		value	of life insurance
		value theory	absolute value		value	of assets		value	of the insolvency
		value changes	firm value	customer	value	approach		value	of the portfolio
		value increases	mean value	same market	value			value	of the property
		value maximization	[1] face value		value	of life		value	of the stockholders
		value reducing	property value	increase the	value		present	value	of aggregate
			book value	expected present	value		face	value	of life
			company value	extreme	value	theory	increase in the	value	
			[2] surrender value		value	and flexible	positive net present	value	
			equity value	optimal face	value			value	and flexible distributions
			total value		value	of aggregate		value	of aggregate loss
			account value		value	of human		value	of human capital
			criterion value		value	of loss		value	of the assets
			shareholder value		value	of pension		value	of the bonds
			asset value		value	of unity		value	of the claims
			critical value		value	of utility		value	of the future
			embedded value		value	of housing		value	of the insurance
			median value					value	of the insurer
			average value					value	of the loss
			customer value						
			discounted value						

[1] *face value*: the amount of life insurance policy.
[2] *surrender value*: the amount of money a policyholder would receive when the policyholder voluntarily cancel the policy before it maturity.

14	coverage	coverage insurance	insurance coverage	health insurance	coverage		private health insurance	coverage	
		coverage correlation	full coverage		coverage	risk correlation	level of insurance	coverage	
		coverage limit	loss coverage	[1] full	coverage	insurance	choice of insurance	coverage	
		coverage risk	partial coverage	medical insurance	coverage		increase in loss	coverage	
		coverage levels	damage coverage	physical damage	coverage		purchase full	coverage	insurance
		coverage rate	universal coverage	risk	coverage	correlation	auto physical damage	coverage	
		coverage contract	reinsurance coverage	rates of	coverage		improvement in loss	coverage	
		coverage period	risk coverage	purchase full	coverage		insurance	coverage	and claim
		coverage offered	liability coverage	obtain full	coverage		availability of insurance	coverage	
		coverage policy	generous coverage	unit of	coverage		consumers obtain full	coverage	
		coverage provided	group coverage	partial	coverage	insurance	contract with full	coverage	
		coverage probability	health coverage	level of	coverage			coverage	and claim occurrence
		coverage ratio	comprehensive coverage	life insurance	coverage			coverage	in monetary terms

No.	Word	2-Word Clusters		3-Word Clusters			4-Word Clusters		
		on the right	on the left						
		coverage against	offer coverage	[2] types of	coverage		health insurance	coverage	rate
		coverage costs	terrorism coverage	availability of	coverage		higher rates of	coverage	
		coverage effects	complete coverage	buy partial	coverage		insurance	coverage	and health
		coverage pooling	provide coverage	choice of	coverage		optimality of full	coverage	
		coverage purchased	flood coverage	insurance	coverage	rate	partial	coverage	in monetary
			building coverage	amount of	coverage		purchase partial	coverage	insurance
			conditional coverage	cost of	coverage				
			employer coverage		coverage	and claim			
			market coverage		coverage	and health			
			maximum coverage	employer sponsored	coverage				
			family coverage	increase loss	coverage				
			public coverage						
	[1] *Full coverage insurance*: the coverage that is paid to the insurer up to the policy limit without subtraction of a deductible.								
	[2] In this study, types of insurance coverage are found in clusters such as *damage coverage, reinsurance coverage, liability coverage, group coverage, health coverage, terrorism coverage, flood coverage, private health insurance coverage*, and *auto physical damage coverage*.								
15	variable	[1] variable annuities	dependent variable		variable	of interest			
		variable measuring	random variable	binary	variable	equal			
		variable life	indicator variable	embedded in	variable				
			binary variable	elliptical random	variable				
			control variable	flood risk	variable				
			explanatory variable	hedging	variable	annuities			
			independent variable	interest rate	variable				
			state variable						
			choice variable						
			conditioning variable						
			risk variable						
			hedging variable						
			ownership variable						
			rate variable						
	[1] *variable annuities*: an insurance product which provides a long term tax deferred savings plan.								
16	policy	[1] policy period	insurance policy	life insurance	policy		increase in	policy	reserve
		policy year	investment policy	increase in	policy		inception of the	policy	
		policy months	public policy	health insurance	policy		value of the	policy	
		policy implications	risk policy	insurers investment	policy				
		policy limit	deductible policy	low risk	policy				
		policy reserve	contribution policy	public	policy	perspective			
		policy variables	optimal policy	pure deductible	policy				
		policy perspective	audit policy	type of	policy				
		policy changes	dividend policy	[2] employer contribution	policy				

No.	Word	2-Word Clusters		3-Word Clusters			4-Word Clusters		
		on the right	on the left						
		policy depends	health policy	outlier payment	policy				
		policy discussions	life policy	partial coverage	policy				
		policy duration	coverage policy						
		policy restrictions	management policy						
		policy types	monetary policy						
		policy form	important policy						
		policy offers	outlier policy						
		policy options	payment policy						
		policy premiums	pricing policy						
		policy strategy	single policy						
			allocation policy						
			premium policy						
			state policy						

[1] *policy period*: the exact duration of the policy with date of when the coverage begins and expires.
[2] *employer contribution policy*: the involvement in providing health insurance options the employers have on their employees.

17	claim	claim effect	aggregated claim		claim	count growth	direct defense and	claim	
		claim count	expected claim		claim	settlement costs	insurance coverage and	claim	
		[1] claim deductible	lower claim	defense and	claim		coverage and	claim	occurrence
		claim payment	suspicious claim		claim	deductible system	defense and	claim	containment
		claim settlement	equity claim	coverage and	claim				
		claim occurrence	higher claim		claim	containment costs			
		claim behaviour	stockholders claim		claim	processing time			
		claim amount	average claim	proof of	claim				
		claim coefficient	contingent claim						
		claim size							
		claim containment							
		claim information							
		claim distribution							
		claim frequency							
		claim tendency							
		claim value							
		claim characteristics							
		claim cost							
		claim patterns							
		claim processing							

[1] *claim deductible*: the amount of money a policyholder needs to pay per claim.

18	health	health insurance	private health	private	health	insurance	demand for	health	insurance
		health care	employer health		health	insurance coverage	demand for private	health	
		health status	life health		health	care spending	private	health	insurance coverage

No.	Word	2-Word Clusters		3-Word Clusters			4-Word Clusters			
		on the right	on the left							
		health insurers	mental health		health	insurance articles	efficiency of	health	production	
		health production	provided health	demand for	health		percentage of	health	insurance	
		health services	financial health	employer	health	insurance	state of the	health		
		health benefits	ill health		health	care reform	completeness of	health	insurance	
		health plan	individual health		health	care system	enrolment in private	health		
		health economy	national health		health	insurance market	prevalence of	health	insurance	
		health coverage	poor health	[1] accident and	health		choice of	health	insurance	
		health risk	public health		health	care costs	cost of	health	care	
		health savings	universal health	provided	health	insurance	employer sponsored	health	insurance	
		health insurer	sponsored health	choice of	health		form of	health	insurance	
		health policy	future health	efficiency of	health		population without	health	insurance	
		health inputs	good health		health	production efficiency	preferences for	health	insurance	
		health economists	receive health	offer	health	insurance	competitive	health	insurance market	
		health maintenance	better health	percentage of	health		employer	health	insurance mandates	
		health reform	group health		health	insurance markets		health	care and insurance	
		health expenditures	mandated health	mental	health	services		health	insurance coverage rate	
		health outcomes	subsidized health	completeness of	health		insurance coverage and	health		
		health spending	underlying health		health	insurance benefits	life and	health	insurance	
		health business	appropriate health	cost of	health		price of	health	insurance	
		health lines	average health		health	insurance plan	private	health	insurance plans	
		health related	competitive health		health	insurance underwriting	self reported	health	status	

[1] *accident and health*: an insurance product which provides coverage for the expenses incurred from an accidental injury, an accidental death, or sickness.

19	individual	individual market	average individual		individual	and group	people in the	individual	
		individual mandate	averse individual		individual	market sample		individual	in age group
		individual insurer	security individual	risk averse	individual		social security	individual	accounts
		individual insurance	single individual	security	individual	accounts		individual	and group market
		individual risk	potential individual		individual	health insurance		individual	and small group
		individual health			individual	in age		individual	of risk class
		individual level		social security	individual				
		individual purchases		increases the	individual				
		individual accounts			individual	consumes less			
		individual prevention			individual	life insurance			
		individual consumes			individual	of risk			
		individual firms							
		individual medical							
		individual annuity							
		individual characteristics							
		individual countries							
		individual firm							
		individual life							

No.	Word	2-Word Clusters		3-Word Clusters			4-Word Clusters		
		on the right	on the left						
		individual losses							
		individual policies							
		individual company							
		individual coverage							
		individual demand							
		individual members							
20	level	level data	optimal level	level	of insurance		optimal	level	of insurance
		level growth	wealth level	level	of risk			level	of insurance coverage
		level demographic	utility level	level	of care			level	of risk aversion
		level characteristics	safety level	level	of quality			level	of equity capital
		level consumption	confidence level	level	of financial			level	of academic dishonesty
		level variation	higher level	level	of market			level	of first party
		level premium	premium level	level	of utility		depends on the	level	
			risk level	level	of education		exceeds a certain	level	
			state level	level	of coverage			level	of additional satisfaction
			certain level	level	of equity			level	of financial development
			individual level	level	of wealth			level	of managerial discretion
			effort level	level	of consumption		optimal	level	of prevention
			reserve level	level	of feedback		same	level	of risk
			critical level	level	of liberalization				
			deductible level	level	of prevention				
			lower level	level	demographic characteristics				
			significance level	level	of academic				
			volatility level	level	of additional				
			education level	level	of capital				
			group level	level	of control				
			state level	level	of effort				
			company level	level	of losses				
			increasing level	level	of satisfaction				
			minimum level	level	of significance				
			overall reserve	level					
21	contract	contract choice	insurance contract		contract	choice effect	inception of the	contract	
		contract offered	pooling contract	fixed price	contract			contract	for the patient
		contract designed	optimal contract	first best	contract		insurance	contract	choice effect
		contract space	new contract	optimal reinsurance	contract		risk first best	contract	
		contract form	reinsurance contract	single	contract	offers		contract	with full coverage
			equilibrium contract	insurance	contract	choice effect			
			single contract	life insurance	contract				
			price contract	low risk	contract				
			best contract	optimal insurance	contract				
			coverage contract	optimal pooling	contract				

No.	Word	2-Word Clusters		3-Word Clusters			4-Word Clusters		
		on the right	on the left						
			risk contract	private insurance	contract				
			separating contract	full coverage	contract				
			buy contract	keep the	contract				
			fair contract	offering a	contract				
			deductible contract	type of	contract				
			deviating contract						
			prefer contract						
			sharing contract						
22	company	company value	insurance company	[2] mutual insurance	company		value of the	company	
		company group	stock company	insurance	company	invests			
		company operations	holding company	insurance	company	operations			
		company invests	mutual company	citizens insurance	company				
		company level	parent company	[3] limited liability	company				
		company executives	[1] ceding company						
			reinsurance company						
			individual company						
			liability company						

[1] *ceding company*: a primary insurance company that transfer, all or part the risk, to other insurers or reinsurers.
[2] *mutual insurance company*: a company in which policyholders are the owners of the company.
[3] *limited liability company*: a company in which owners have limited personal liability for the debts.

23	capital	capital market	equity capital	cost of	capital		proportion of human	capital	
		capital requirement	[1] human capital	[2] risk based	capital			capital	asset pricing model
		capital structure	solvency capital	[3] solvency	capital	requirement	amount of equity	capital	
		capital allocation	based capital		capital	asset pricing	cost of equity	capital	
		capital budgeting	internal capital	amount of	capital		costs of internal	capital	
		capital ratio	required capital		capital	budgeting decisions	level of equity	capital	
		capital support	external capital		capital	market line	funds and venture	capital	
		capital asset	risk capital	human	capital	insured	important for	capital	asset
		capital adequacy	economic capital		capital	structure decisions	risk on	capital	structure
		capital insured	optimal capital		capital	support risk	risk based	capital	formula
		capital standards	minimum capital	minimum required	capital		value of human	capital	
		capital required	regulatory capital		capital	allocation method	volume of	capital	flowing
		capital cost	venture capital		capital	market theory			
		capital providers	hoarding capital	amounts of	capital				
		capital gains	target capital		capital	market participants			
		capital raised	debt capital		capital	to asset			
		capital charge	safety capital	optimal	capital	requirement			
		capital held	available capital	supply of	capital				
		capital increases	bank capital		capital	structure studies			
		capital models	financial capital	effects on	capital				

No.	Word	2-Word Clusters		3-Word Clusters			4-Word Clusters		
		on the right	on the left						
		capital needed	insurer capital	level of	capital				
		capital shock	additional capital	approach to	capital				
		capital investment	holding capital	based	capital	formula			
		capital management	insurance capital						
		capital base							
	[1] *human capital*: as employees in an organisation, human have been seen as a capital investment, such that education or trainings, would improve their performance (in terms of quality and level of production)								
	[2] *risk based capital*: a method which is used to find an amount of capital that an insurance company should have to support the business operations.								
	[3] *solvency capital requirement*: an insurance regulation which concerns the amount of capital an insurance company is required to hold as to prevent the risk of lack of financial liquidity.								
24	measure	measure firm	risk measure		measure	of risk			
		measure insurance	probability measure	risk adjusted	measure				
		measure provides	adjusted measure	entropic risk	measure				
		measure risk	performance measure		measure	of insurance			
			elasticity measure		measure	of market			
			size measure		measure	of firm			
			neutral measure	risk neutral	measure				
			pricing measure	entropy probability	measure				
			objective measure		measure	firm size			
					measure	the level			
25	financial	financial distress	insurer financial	poor	financial	practices	likelihood of	financial	distress
		financial institutions	dynamic financial	dynamic	financial	analysis	number of	financial	analysts
		financial services	poor financial		financial	market returns	differences in poor	financial	
		financial market	good financial	level of	financial		insurers with higher	financial	
		financial crisis	international financial	good	financial	practices	level of	financial	development
		financial strength	greater financial	likelihood of	financial		level of	financial	expertise
		financial practices	nonbank financial	accounting and	financial				
		financial assets	global financial	differences in	financial				
		financial expertise	optimal financial		financial	services market			
		financial firms	significant financial	nonbank	financial	institutions			
		financial wealth	consolidated financial	types of	financial				
		financial statements	higher financial		financial	and actuarial			
		financial reporting	overall financial		financial	decision making			
		financial performance	personal financial		financial	services firms			
		financial risk	total financial	costs of	financial				
		financial analysis		economic and	financial				
		financial health			financial	and insurance			
		financial instruments			financial	practices index			
		financial loss			financial	reporting standards			
		financial analysts			financial	risk management			

No.	Word	2-Word Clusters		3-Word Clusters			4-Word Clusters		
		on the right	on the left						
		financial development			financial	services industry			
		financial intermediaries			financial	services sector			
		financial leverage							
		financial decision							
		financial quality							
26	year	year period	[2] policy year	percent per	year			year	developed loss ratios
		year end	every year		year	to year			
		[1] year floodplain	first year		year	fixed effects			
		year developed	given year	end of	year				
		year fixed	current year		year	developed loss			
		year term	fiscal year		year	policy period			
		year effects	last year		year	loss reserve			
		year flood	next year		year	term life			
		year time	current year						
		year olds	previous year						
		year ahead	accident year						
		year based	ending year						
		year dummies	calendar year						
		year observations	prior year						
		year change	same year						
		year earnings	following year						
		year preceding	sample year						
			single year						
	[1] *policy year*: a12-month period of the effective date of insurance coverage beginning on the date the policy starts and ending on the same date of the following year.								
	[2] *year floodplain*: or the full term 100-year floodplain. It is the method used in calculating the probability of flooding in a predicted area. The flood insurance rate will correspond to the 100-year floodplain.								
27	case	case scenario	special case	consider the	case		risk management	case	project
		case project	general case	worst	case	scenario			
		case study	latter case	adverse selection	case				
			worst case	best	case	scenario			
			reference case		case	of insurance			
			first case	management	case	project			
			baseline case	risk management	case				
			best case						
			extreme case						
			interesting case						
			limiting case						
			premium case						
			second case						
			selection case						
			management case						

No.	Word	2-Word Clusters		3-Word Clusters			4-Word Clusters		
		on the right	on the left						
			new case						
28	period	period model	sample period		period	of time	end of the	period	
		period return	probationary period		period	cash flow	long	period	of time
		period assets	year period	end of	period		first	period	cash flow
		period framework	study period	first	period	cash	end of	period	assets
		period approach	policy period	pure probationary	period			period	and limited term
		period compared	first period	[1] holding	period	return	probationary	period	and limited
		period consumption	single period		period	and limited	beginning of the	period	
		period depending	holding period	year	period	period	combination of probationary	period	
		period lead	reference period	bull market	period		losses incurred in	period	
		period problem	coverage period	entire sample	period				
			deferred period						
			estimation period						
			initial period						
			last period						
			future period						
			information period						
			market period						
			prior period						
			entire period						
			extended period						
			next period						
			long period						
			previous period						

[1] *holding period return*: a measure used in evaluating the performance of investment. It looks at the total return on assets over the period which those assets have been holding.

29	probability	probability distribution	survival probability		probability	of loss		probability	of a loss
		probability measure	shortfall probability		probability	of death		probability	of an accident
		probability density	loss probability		probability	density function	minimal entropy	probability	measure
		probability distortion	fraud probability		probability	of incurring	risks	probability	of incurring
		probability increases	high probability		probability	distortion functions			
		probability curve	conditional probability		probability	of shortfall			
		probability level	prior probability	reduce the	probability				
		probability decreases	accident probability		probability	of fraud			
		probability estimates	default probability		probability	of bankruptcy			
		probability function	low probability		probability	of occurrence			
		probability index	average probability		probability	of survival			
		probability mass	bankruptcy probability	increase the	probability				
		probability model	coverage probability	minimal entropy	probability				
			entropy probability		probability	density functions			
			insolvency probability		probability	of default			

No.	Word	2-Word Clusters		3-Word Clusters			4-Word Clusters		
		on the right	on the left						
			neutral probability	shortfall	probability	curve			
			posterior probability	entropy	probability	measure			
			annual probability	low risks	probability				
			base probability		probability	of voluntary			
			greater probability	unconditional fraud	probability				
			joint probability						
			linear probability						
			marginal probability						
			risks probability						
			surrender probability						
30	life	life insurance	non life	demand for	life		demand for	life	insurance
		life insurer	term life		life	insurance demand	amount of	life	insurance
		life expectancy	whole life		life	insurance market		life	insurance and annuity
		[1] life annuity	real life		life	insurance products		life	insurer stock returns
		life health	international life		life	insurance industry		life	insurance and stock
		life cycle	subjective life		life	insurance contracts	value of	life	insurance
		life table	[2] universal life		life	insurance purchase	determinants of	life	insurance
		life settlement	participating life		life	health insurers		life	and non life
		life market	purchase life	amount of	life		optimal amount of	life	
		life policy	single life		life	insurance company	options in	life	insurance
		life business	mutual life		life	insurance consumption	rate sensitivity of	life	
		life premiums	average life	[3] term	life	insurance	aggregate	life	insurance demand
		life reinsurance	individual life	international	life	insurance	consumption and	life	insurance
			private life		life	and annuity	duration of	life	insurance
			stock life		life	insurance policy	face value of	life	
			working life	non	life	insurance		life	and annuity insurance
			aggregate life	participating	life	insurance		life	and health insurance
			buy life		life	and health	participating	life	insurance contracts
			global life		life	insurance coverage	percent of the	life	
			large life		life	insurance premiums	predictions of	life	tables
			domestic life		life	insurer stock	real	life	insurance demand
			entire life	universal	life	insurance	returns of	life	insurers
			group life	value of	life				
			longer life						
			optimal life						

[1] *life annuity*: a type of an insurance product which guarantee a periodic income payment to the annuitants throughout the life of the annuitants named in the contract.
[2] *universal life*: a type of life insurance product which an insured could adjust the amount of premium payment. The product is design to have flexible features with cash accumulated value.
[3] *term life*: unlike whole life insurance which provides life coverage until the death of an insured, the term life insurance covers the insured person only within a specified period such as 10 or 20 years.

31	return	return distribution	stock return	rate of	return		free rate of	return	
		return volatility	expected return	risk and	return		expected rate of	return	

No.	Word	2-Word Clusters		3-Word Clusters			4-Word Clusters		
		on the right	on the left						
		return series	abnormal return		return	on assets		return	on invested assets
		return shift	asset return	fair expected	return		real rate of	return	
		return profile	investment return	holding period	return			return	on the market
		return data	log return	long	return	series	expected	return	and volatility
		return premium	fair return		return	and volatility	fair rate of	return	
		return problem	period return	mean	return	shift	rate of	return	targeting
		return risk	total return		return	on investment		return	on their investment
		return targeting	adjusted return	cumulative abnormal	return		total rates of	return	
			excess return	fair	return	premium			
			mean return		return	on capital			
			risk return		return	per period			
			average return						
			equity return						
32	liability	[1] liability insurance	property liability	property	liability	insurance	property	liability	insurance industry
		liability claims	asset liability	asset	liability	management	joint and several	liability	
		liability management	auto liability	property and	liability		non property	liability	insurers
		liability coverage	limited liability	non property	liability		demand for	liability	insurance
		liability ratio	environmental liability		liability	insurance industry	private passenger auto	liability	
		liability insurer	general liability	asset and	liability		property and	liability	insurance
		liability lines	automobile liability	demand for	liability		property and liability	liability	insurers
		liability provision	party liability		liability	insurance companies	value to	liability	ratio
		liability matching	commercial liability	property	liability	insurer			
		liability cash	legal liability	[2] bodily injury	liability				
		liability exposure	tort liability	asset	liability	matching			
		liability loss	tax liability		liability	cash flows			
		liability reforms	medical liability	limited	liability	provision			
		liability company	several liability	long term	liability				
		liability cover	injury liability	medical	liability	insurance			
		liability crisis	term liability	property	liability	lines			
		liability market	catastrophic liability	auto	liability	insurance			
		liability proportion	malpractice liability	automobile	liability	claims			
		liability risk	constant liability	availability of	liability				
			total liability		liability	and life			
				limited	liability	company			
				passenger auto	liability				
				[3] third party	liability				

[1] *liability insurance*: a type of general insurance product which provide the coverage of loss from being sued by third parties for negligence.
[2] *bodily injury liability*: an insurance product designed to provide legal liability coverage in the event of an accident to passengers in the insured's car and other cars.
[3] *third party liability*: a type of insurance covering the policyholders in the case of the legal obligation of third parties.

33	reinsurance	reinsurance market	optimal reinsurance	insurance and	reinsurance		insurance and	reinsurance	markets

No.	Word	2-Word Clusters		3-Word Clusters			4-Word Clusters		
		on the right	on the left						
		reinsurance industry	catastrophe reinsurance		reinsurance	and securitization	insurance and	reinsurance	companies
		reinsurance assets	[1] terrorism reinsurance	amount of	reinsurance		holdings of	reinsurance	assets
		reinsurance coverage	private reinsurance	demand for	reinsurance			reinsurance	assumed from affiliates
		reinsurance contract	global reinsurance	holdings of	reinsurance		small holdings of	reinsurance	
		reinsurance decisions	traditional reinsurance	net of	reinsurance				
		reinsurance program	government reinsurance	insurance or	reinsurance				
		reinsurance variables	federal reinsurance	price of	reinsurance				
		reinsurance marketplace	foreign reinsurance	optimal	reinsurance	contract			
		reinsurance premium	purchase reinsurance		reinsurance	for terrorism			
		reinsurance company	world reinsurance	catastrophe	reinsurance	market			
		reinsurance expense	conventional reinsurance	credit for	reinsurance				
		reinsurance policies	international reinsurance	private terrorism	reinsurance				
		reinsurance costs	life reinsurance	property and	reinsurance				
		reinsurance directive	replace reinsurance	securitization and	reinsurance				
		reinsurance receivables							
		reinsurance strategy							
		reinsurance usage							
		reinsurance transaction							

[1] *terrorism reinsurance*: a reinsurance product, which was introduced after the events of September 11, 2001, designed to provide coverage protection against property losses that might occur from terrorist attacks.

34	function	function based	utility function	cumulative distribution	function		absolute risk aversion	function	
			distribution function	power utility	function				
			cost function	probability density	function				
			density function	risk aversion	function				
			objective function	square root	function				
			value function	expected utility	function				
			production function	health production	function				
			distortion function	morgenstern utility	function				
			increasing function	concave utility	function				
			characteristic function		function	of firm			
			demand function	exponential utility	function				
			loss function	inverse demand	function				
			aversion function	natural logarithm	function				
			indicator function						
			linear function						
			root function						
			differentiable function						
			power function						
			activation function						
			criterion function						

No.	Word	2-Word Clusters		3-Word Clusters			4-Word Clusters		
		on the right	on the left						
			logistic function						
			decreasing function						
			intermediation function						
			logarithm function						
35	analysis	analysis reveals	empirical analysis	dynamic financial	analysis		included in the	analysis	
		analysis shows	regression analysis	cross sectional	analysis		focus of the	analysis	
		analysis indicates	financial analysis	panel data	analysis				
		analysis provides	numerical analysis		analysis	of insurance			
		analysis focuses	sensitivity analysis	cross frontier	analysis				
		analysis includes	detailed analysis	multiple regression	analysis				
		analysis suggests	factor analysis						
			statistical analysis						
			economic analysis						
			frontier analysis						
			cluster analysis						
			dynamic analysis						
			multivariate analysis						
			sectional analysis						
			welfare analysis						
			data analysis						
			run analysis						
			further analysis						
			scenario analysis						
			series analysis						
			credit analysis						
			equilibrium analysis						
			risk analysis						
36	utility	utility function	expected utility	diminishing marginal	utility		aversion and	utility	theory
		utility theory	marginal utility	loss of	utility		descriptions of	utility	theory
		utility level	power utility	level of	utility		risk aversion and	utility	
		utility curve	additional utility	power	utility	function		utility	theory and risk
		utility associated	exponential utility		utility	without insurance	expected	utility	without insurance
		utility maximization	concave utility	expected	utility	maximization	marginal	utility	of wealth
		utility equal	corresponding utility	negative exponential	utility			utility	with no insurance
		utility derived	elastic utility		utility	of wealth	amount of additional	utility	
		utility generated	lifetime utility	aversion and	utility		expected	utility	with insurance
		utility maximizing	morgenstern utility	descriptions of	utility		introduction to	utility	theory
		utility maximizers	higher utility	expected	utility	function		utility	derived from wealth
		utility evaluation	indirect utility	morgenstern	utility	function		utility	with insurance priced
		utility cost	logarithmic utility		utility	by buying			
		utility creates	same utility	calculate the	utility				

No.	Word	2-Word Clusters		3-Word Clusters			4-Word Clusters		
		on the right	on the left						
		utility loss	aversion utility	concave	utility	function			
		utility terms		expected	utility	maximizers			
		utility value		exponential	utility	function			
				maximize expected	utility				
					utility	of consumption			
37	optimal	optimal asset	socially optimal		[1] optimal	asset allocation		optimal	product mix proportion
		optimal level	corresponding optimal		optimal	product mix		optimal	level of insurance
		optimal product		calculate the	optimal			optimal	amount of life
		optimal contract		compute the	optimal		calculate the	optimal	product
		optimal reinsurance		derive the	optimal			optimal	level of prevention
		optimal risk			optimal	risk sharing			
		optimal consumption		find the	optimal				
		optimal insurance			optimal	capital requirement			
		optimal amount			optimal	private insurance			
		optimal solution			optimal	reinsurance contract			
		optimal decisions		determine the	optimal				
		optimal capital			optimal	face value			
		optimal hedging			optimal	insurance contract			
		optimal choice			optimal	pooling contract			
		optimal control			optimal	tax deduction			
		optimal policy			optimal	for policyholders			
		optimal stock			optimal	hedging strategy			
		optimal investment			optimal	risk management			
		optimal strategy			optimal	solution set			
		optimal tax			optimal	solvency capital			
		optimal values			optimal	stock purchase			
		optimal demand		present the	optimal				
		optimal face		theory of	optimal				
		optimal indemnity							
	[1] *optimal asset allocation*: a concept that aims for maximum return on investment.								
38	ratio	ratio increases	loss ratio	assets inputs	ratio		consumption to wealth	ratio	
		ratio dynamics	dependency ratio	dependency	ratio	dynamics	debt to gdp	ratio	
		ratio falls	wealth ratio	incremental reserve	ratio			ratio	dynamics for population
		ratio volatility	combined ratio	mean dependency	ratio		value to liability	ratio	
			capital ratio	dependency	ratio	volatility			
			gdp ratio	income replacement	ratio				
			cession ratio	target buffer	ratio				
			expense ratio						
			liability ratio						
			likelihood ratio						

No.	Word	2-Word Clusters		3-Word Clusters			4-Word Clusters		
		on the right	on the left						
			funding ratio						
			reserve ratio						
			replacement ratio						
			buffer ratio						
			surplus ratio						
			assets ratio						
			equity ratio						
			inputs ratio						
			coverage ratio						
			hedge ratio						
			leverage ratio						
			median ratio						
			premium ratio						
39	plan	plan quality	public plan	average	plan	generosity	offering a single	plan	
		plan generosity	insurance plan	risk management	plan		choice of insurance	plan	
		plan choice	health plan	health insurance	plan		long term incentive	plan	
		plan characteristics	average plan	health	plan	characteristics	number of	plan	restrictions
		plan premiums	pension plan	health	plan	choice			
		plan design	management plan		plan	fixed effects			
		plan features	single plan	pension	plan	participants			
		plan participants	pricing plan	term incentive	plan				
		plan type	withdrawal plan						
		plan fixed	care plan						
		plan restrictions	incentive plan						
			retirement plan						
40	demand	demand curve	insurance demand		demand	for insurance		demand	for life insurance
		demand elasticity	annuity demand		demand	for life	price elasticity of	demand	
		demand patterns	care demand	life insurance	demand			demand	for medical care
		demand function	asset demand	elasticity of	demand			demand	for health insurance
		demand insurance	aggregate demand		demand	for medical		demand	for private health
		demand increases	hedging demand		demand	for private	aggregate	demand	for private
		demand model	increased demand	medical care	demand		related to the	demand	
		demand reaction	excess demand		demand	for health	change in asset	demand	
		demand equation	optimal demand		demand	for reinsurance	aggregate life	demand	demand
		demand response	consumer demand		demand	for annuities	assuming a	demand	elasticity
		demand variables	higher demand	supply and	demand		corporate	demand	for insurance
			lower demand	affect the	demand			demand	for insurance products
			group demand		demand	for liability		demand	for liability insurance
			inverse demand	change in	demand		impact of asset	demand	
			corporate demand	annuity	demand	reaction			
			greater demand	increase the	demand				

No.	Word	2-Word Clusters		3-Word Clusters			4-Word Clusters		
		on the right	on the left						
			high demand	insurance	demand	increases			
			individual demand	low annuity	demand				
				assuming a	demand				
				care	demand	curve			
					demand	for assets			
					demand	for auto			
					demand	for coverage			
				health insurance	demand				
				hedging	demand	model			
41	distribution	distribution function	normal distribution	cumulative	distribution	function		distribution	of the random
		distribution channel	stable distribution	standard normal	distribution		follow the normal	distribution	
		distribution system	loss distribution	alternative	distribution	functions			
		distribution corresponds	probability distribution		distribution	of quality			
		distribution mechanism	elliptical distribution	multivariate normal	distribution				
		distribution set	cumulative distribution	stable	distribution	set			
			joint distribution						
			return distribution						
			surplus distribution						
			broker distribution						
			marginal distribution						
			alternative distribution						
			claim distribution						
			conditional distribution						
			income distribution						
			insurance distribution						
			prior distribution						
			severity distribution						
			size distribution						
			value distribution						
			flow distribution						
			generalized distribution						
42	purchase	purchase insurance	insurance purchase		purchase	of insurance	general risk policyholders	purchase	
		purchase full	policyholders purchase	risk policyholders	purchase		decision to	purchase	insurance
		purchase partial	annuity purchase	life insurance	purchase			purchase	full coverage insurance
		purchase annuities	stock purchase	decision to	purchase		specific risk policyholders	purchase	
		purchase life	mandatory purchase		purchase	more insurance	policyholders	purchase	full insurance
		purchase requirement	individuals purchase		purchase	no insurance	policyholders	purchase	partial insurance
		purchase decision	optimally purchase		purchase	full coverage		purchase	partial coverage insurance
		purchase decisions		mandatory	purchase	requirement			
		purchase reinsurance			purchase	full insurance			
		purchase actuarially			purchase	life insurance			

No.	Word	2-Word Clusters		3-Word Clusters			4-Word Clusters		
		on the right	on the left						
		purchase coverage		insurance	purchase	decision			
		purchase health		policyholders	purchase	insurance			
					purchase	from insurers			
					purchase	partial insurance			
				require the	purchase				
				incentive to	purchase				
				income individuals	purchase				
				optimal stock	purchase				
				prefer to	purchase				
					purchase	partial coverage			
43	wealth	wealth ratio	initial wealth	root of	wealth		marginal utility of	wealth	
		wealth level	preannuitized wealth	income and	wealth		proportion of preannuitized	wealth	
		wealth increases	financial wealth	level of	wealth			wealth	and future income
		wealth invested	total wealth	utility of	wealth		remaining	wealth	after paying
		wealth accumulation	future wealth		wealth	and future	utility derived from	wealth	
		wealth effect	expected wealth	expected future	wealth			wealth	and asset price
		wealth equals	shareholder wealth	increase in	wealth			wealth	in the loss
		wealth equivalent	final wealth		wealth	and asset			
		wealth transfer	pension wealth	change in	wealth				
			annuity wealth	consumption and	wealth				
			remaining wealth	derived from	wealth				
			current wealth		wealth	after paying			
			housing wealth						
			marketable wealth						
			positive wealth						
			adjusted wealth						
			different wealth						
			household wealth						
			national wealth						
			nonhousing wealth						
			optimal wealth						
			personal wealth						
			terminal wealth						
44	information	information asymmetry	private information	asymmetric	information	problems	disclosure of quality	information	
		information available	asymmetric information		information	about policyholders	existence of asymmetric	information	
		information set	public information		information	about risk		information	
		information regarding	superior information	source of	information		insurers that publicize	information	
		information quality	quality information	publicize their	information		based on the	information	
		information sharing	publicize information	publicizing of	information				

No.	Word	2-Word Clusters		3-Word Clusters			4-Word Clusters		
		on the right	on the left						
		information costs	imperfect information	receives private	information				
		information becomes	provides information	impact on	information				
		information disclosure	full information						
		information content	genetic information						
		information period	available information						
		information systems	additional information						
		information advantage	claim information						
		information problem	accounting information						
		information revealed	hidden information						
		information exists	important information						
		information sheet	insurance information						
		information system	obtain information						
			provide information						
			publicizing information						
			regulatory information						
			risk information						
			descriptive information						
			financial information						
			personal information						
45	factor	factor analysis	loading factor	asset risk	factor		impatience	factor	for consumption
		factor model	risk factor		factor	for consumption		factor	for future consumption
			discount factor	proportional loading	factor				
			impatience factor		factor	for future			
			important factor	product risk	factor				
			adjustment factor	risk aversion	factor				
			systematic factor						
			aversion factor						
			major factor						
			significant factor						
46	hedge	hedge fund	optimal hedge	optimal	hedge	ratios	types of managers	hedge	
		hedge portfolio	managers hedge		hedge	fund division			
		hedge ratio	estimated hedge						
47	management	management action	risk management		management	and insurance	risk	management	and insurance
		management activities	liability management	corporate risk	management			management	and insurance research
		management process	exposure management	risk	management	action		management	and insurance journals
		management control	utilization management	asset liability	management		cost of risk	management	
		management decision	asset management	risk	management	activities		management	and insurance program
		management plan	capital management	sustainability risk	management		articles in risk	management	
		management behaviour	claims management	internal risk	management			management	and insurance major
		management policies	effective management	risk	management	process		management	of catastrophe risk

No.	Word	2-Word Clusters		3-Word Clusters			4-Word Clusters		
		on the right	on the left						
		management skills	efficient management	enterprise risk	management		risk	management	case project
		management issues	passive management	risk	management	strategy			
		management perspective	insurer management	risk	management	plan			
		management policy	strategic management	leading risk	management				
		management strategy		environmental risk	management				
		management techniques		risk	management	decision			
		management cost			management	and control			
		management department			management	of catastrophe			
		management measures		risk	management	cost			
		management case		risk	management	issues			
		management concepts		risk	management	measures			
		management mix		voluntary risk	management				
		management models		catastrophe risk	management				
		management program		commercial risk	management				
				comprehensive risk	management				
48	investment	investment policy	insurers investment		investment	in prevention	return on their	investment	
		investment strategy	net investment	net	investment	income			
		investment income	risky investment	static	investment	strategy			
		investment risk	optimal investment	underwriting and	investment				
		investment return	free investment	insurers	investment	policy			
		investment centre	initial investment	nonstatic	investment	strategy			
		investment portfolio	nonstatic investment	reregulating the	investment				
		investment strategies	static investment	return on	investment				
		investment decisions	capital investment	risk free	investment				
		investment grade	direct investment	cost of	investment				
		investment opportunities	financial investment	foreign direct	investment				
		investment rate	additional investment		investment	in stocks			
		investment products	lower investment	level of	investment				
		investment banking		profitability and	investment				
		investment guarantees							
		investment model							
		investment performance							
48	annuity	annuity demand	life annuity	insurance and	annuity		life insurance and	annuity	
		annuity products	variable annuity	life and	annuity		duration of the	annuity	
		annuity purchases	effective annuity		annuity	demand reaction	adverse selection on	annuity	
		annuity contracts	individual annuity	life	annuity	contracts	insurance and	annuity	products
		annuity income	voluntary annuity	low	annuity	demand	life and	annuity	insurance
		annuity payouts	fair annuity	[1] single life	annuity				
		annuity wealth	low annuity	level of	annuity				
		annuity market	optimal annuity	selection on	annuity				

No.	Word	2-Word Clusters		3-Word Clusters			4-Word Clusters		
		on the right	on the left						
		annuity premiums							
		annuity business							
		annuity provider							
		annuity payment							
		annuity prices							
		annuity flows							
		annuity insurance							
	[1] *single life annuity*: the two types of annuity: single life annuity and joint life annuity. For a single life annuity, the monthly benefit is only given to the policyholder. On the other hand, a joint life annuity also provide benefit to the person the policyholder choose to share the annuity with.								
50	benefit	benefit restrictions	marginal benefit	pharmacy	benefit	restrictions			
		benefit payments	death benefit	medicaid pharmacy	benefit				
		benefit design	pharmacy benefit						
		benefit plans	defined benefit						
		benefit trade	employee benefit						
		benefit levels	optimal benefit						
		benefit participation	potential benefit						
			diversification benefit						
			pension benefit						
51	stock	stock market	held stock	held	stock	companies	publicly held	stock	companies
		stock price	restricted stock	publicly held	stock		direction of the	stock	
		stock return	owned stock	mutual owned	stock		life insurer	stock	returns
		stock purchase	common stock		stock	market response	life insurance and	stock	
		stock company	insurer stock		stock	risk premium	insurance and	stock	purchases
		stock volatility	subsidiary stock	widely held	stock		mutual and	stock	companies
		stock ownership	optimal stock	mutual and	stock			stock	and mutual insurers
		stock insurer	negative stock		stock	price returns	widely held	stock	companies
		stock risk	preferred stock	insurer	stock	returns			
		stock insurance	firms stock	life insurer	stock				
		stock options	housing stock		stock	insurance companies			
		stock firms		subsidiary	stock	companies			
		stock exchange		common	stock	companies			
		stock fraction		insurance and	stock				
		stock held		owned	stock	companies			
		stock life			stock	ownership classes			
		stock based			stock	price reaction			
		stock awarded			stock	based incentive			
				effect on	stock				
				negative	stock	market			
				optimal	stock	purchase			
				restricted	stock	held			

No.	Word	2-Word Clusters		3-Word Clusters			4-Word Clusters		
		on the right	on the left						
					stock	market performance			
					stock	market reaction			
					stock	price response			
52	term	term care	long term	long	term	care	long	term	care insurance
		term life	short term	long	term	incentive	long	term	incentive plan
		term incentive	error term		term	care insurance	long	term	care costs
		term bonds	limited term	short	term	bonds	period and limited	term	
		term insurance	second term		term	life insurance	end of the	term	
		term interest	first term		term	incentive plan	long	term	interest rates
		term contracts	interaction term	long	term	interest			
		term debt	last term	long	term	debt			
		term investments	year term		term	care costs			
		term disability	residual term	short	term	contracts			
		term liability	adjustment term		term	interest rates			
		term structure	intercept term	long	term	disability			
		term relationship	squared term	long	term	investments			
		term average		long	term	liability			
		term changes		long	term	relationship			
		term growth		long	term	problem			
		term problem		pure limited	term				
		term risk		random error	term				
				short	term	changes			
					term	of internationalization			
				year	term	life			
53	income	income tax	future income		income	and wealth	pre existing pension	income	
		income workers	low income	individuals with	income		wealth and future	income	
		income elasticity	premium income	low	income	workers	price and	income	elasticities
		income individuals	investment income	net investment	income				
		income replacement	pension income	price and	income				
		income security	retirement income	existing pension	income				
		income households	capita income	total premium	income				
		income groups	labor income	government pension	income				
		income levels	higher income		income	and preferences			
		income quintile	annuity income		income	individuals purchase			
		income people	family income		income	replacement ratio			
		income process	wage income		income	tax rate			
		income distribution	taxable income	low	income	individuals			
		income effect	average income						
		income greater	disposable income						
		income risk	higher income						

No.	Word	2-Word Clusters		3-Word Clusters			4-Word Clusters		
		on the right	on the left						
			lower income						
			fixed income						
			household income						
			total income						
			annual income						
			corporate income						
			median income						
			national income						
54	bond	bond market	cat bond	cat	bond	market		bond	covenant for mitigating
		bond covenant	linked bond	cat	bond	spreads		bond	transaction between agent
		bond repayment	promised bond	index linked	bond		mortgage and	bond	default
		bond spreads	corporate bond		bond	repayment amount	promised	bond	repayment amount
		bond prices	coupon bond	cat	bond	issues			
		bond issue	government bond	simple cat	bond				
		bond redemptions		cat	bond	transaction			
		bond transaction		promised	bond	repayment			
		bond insurers		cat	bond	prices			
		bond yields		mortgage and	bond				
		bond default		hybrid cat	bond				
55	property	property liability	residential property		[1]property	casualty insurance		property	and casualty insurance
		property casualty	non property		property	liability insurance		property	liability insurance industry
		property damage	coastal property		property	liability insurers	non	property	liability insurers
		property insurance	commercial property		property	and casualty	value of the	property	
		property value	damaged property		property	damage claims		property	and liability insurance
		property owners	mortgaged property		property	and liability		property	and liability insurers
		property catastrophe	large property	non	property	liability			
		property claims	average property		property	casualty insurers			
		property loss	personal property		property	insurance markets			
		property prices			property	damage losses			
					property	liability lines			
					property	and reinsurance			
					property	claims services			
					property	casualty lines			
	[1] *property casualty insurance*: an insurance product which provides coverage against losses to an insured's property and as well the legal liability that may result from injury or damage to the property of others.								
56	average	average premium	weighted average		average	plan generosity		average	daily trading volume
		average cost	lower average		average	proportional losses		average	number of days
		average number	higher average	higher	average	costs	higher than the	average	
		average plan	greater average	percent on	average				
		average annual	industry average		average	annual growth			

No.	Word	2-Word Clusters		3-Word Clusters			4-Word Clusters		
		on the right	on the left						
		average income	historical average		average	daily trading			
		average loss	moving average		average	life expectancy			
		average value	term average		average	medical expenses			
		average rate			average	premium rate			
		average earnings			average	premium savings			
		average medical		differences in	average				
		average life							
		average price							
		average probability							
		average share							
		average claim							
		average health							
		average mortality							
		average payment							
		average property							
		average return							
		average values							
57	pay	pay claims	employers pay	willing to	pay			pay	for accounting performance
		pay attention	shareholders pay		pay	for insurance		pay	for performance compensation
		pay losses	significant pay		pay	for performance			
				required to	pay				
				ability to	pay				
					pay	a premium			
					pay	for accounting			
				degree of	pay				
				liable to	pay				
					pay	a tax			
					pay	for protection			
					pay	for reconstruction			
				unwilling to	pay				
58	consumer	consumer price	averse consumer		consumer	price index			
		consumer demand	prudent consumer	risk averse	consumer				
		consumer protection							
		consumer heterogeneity							
		consumer risk							
		consumer credit							
		consumer decides							
		consumer response							
		consumer welfare							
59	equilibrium	equilibrium exists	separating equilibrium	long run	equilibrium			equilibrium	in a market
		equilibrium condition	pooling equilibrium	separating	equilibrium	exists	existence of an	equilibrium	

No.	Word	2-Word Clusters		3-Word Clusters			4-Word Clusters		
		on the right	on the left						
		equilibrium value	run equilibrium		equilibrium	pooled premium			
		equilibrium concept	disclosure equilibrium	unique separating	equilibrium				
		equilibrium model	market equilibrium						
		equilibrium premium	partial equilibrium						
		equilibrium analysis	competitive equilibrium						
		equilibrium outcomes	degenerate equilibrium						
		equilibrium pooled	efficiency equilibrium						
		equilibrium price	proposed equilibrium						
60	selection	selection bias	[1] adverse selection	adverse	selection	costs	hazard and adverse	selection	
		selection model	self selection	sided adverse	selection			selection	and moral hazard
		selection component	risk selection	adverse	selection	component	adverse	selection	in insurance
		selection problem	sample selection	self	selection	constraint	adverse	selection	and moral
		selection constraint	favorable selection	adverse	selection	problem	presence of adverse	selection	
		selection incentives	portfolio selection		selection	and moral		selection	in insurance markets
		selection theory	propitious selection		selection	in insurance	insurer sided adverse	selection	
		selection case		adverse	selection	model	double sided adverse	selection	
		selection cost		adverse	selection	case	policyholder sided adverse	selection	
		selection effects		adverse	selection	theory	absence of adverse	selection	
		selection factors		standard adverse	selection		adverse	selection	on annuity
		selection appears		adverse	selection	effects	asymmetry and adverse	selection	
		selection exists		avoid adverse	selection		impact of adverse	selection	
		selection issues			selection	for insurers	markets with adverse	selection	
					selection	of assets	one sided adverse	selection	
					selection	on annuity	problem of adverse	selection	

[1] *adverse selection*: commonly found with many other collocates. It is the tendency of those exposed to a higher risk to seek more insurance coverage than those at a lower risk. Insurers react either by charging higher premiums or not insuring at all, as in the case of floods.

61	rating	rating agencies	disability rating	underwriting and	rating		underwriting and	rating	restrictions
		rating restrictions	experience rating	credit	rating	agencies			
		rating system	credit rating	financial strength	rating				
		rating components	community rating	key	rating	guide			
		rating firms	financial rating						
		rating guide	previous rating						
			current rating						
			firm rating						
			key rating						
			prior rating						
			insurance rating						
			risk rating						
			strength rating						

No.	Word	2-Word Clusters		3-Word Clusters			4-Word Clusters		
		on the right	on the left						
62	group	group market	risk group	[1] individual and	group		individual and	group	markets
		group insurance	small group	higher risk	group		individual in age	group	
		group coverage	age group	lower risk	group		risk and insurance	group	
		group affiliation	insurance group		group	market sample			
		group setting	international group	employment based	group				
		group membership	actuarial group	small	group	market			
		group level	based group		group	health insurance			
		group demand	national group		group	of insurers			
		group health	company group		group	of respondents			
		group oriented	working group	relationship between	group				
		group life	valuation group						
		group plans							
		group size							
	[1] group insurance and individual insurance are types of insurance plans. Individual insurance is designed to provide loss coverage to a single person. On the other hand, group insurance is an insurance plan that provides coverage for two or more individuals.								
63	private	private debt	optimal private		private	health insurance	demand for	private	health
		private health	traditional private		private	debt holdings		private	health insurance coverage
		private information	hold private		private	debt market	aggregate demand for	private	
		private insurance	receives private		private	debt investments	determinants of	private	debt
		private market		enrollment in	private		enrollment in	private	health
		private sector			private	insurance companies	insurers with	private	debt
		private insurer			private	sector workers		private	health insurance plans
		private reinsurance		determinants of	private			[1] private	passenger auto liability
		private passenger		insurers with	private				
		private plans		optimal	private	insurance			
		private life			private	insurance markets			
		private placement		hold	private	debt			
		private terrorism			private	insurance contract			
					private	insurance market			
					private	life insurance			
					private	passenger auto			
					private	terrorism reinsurance			
	[1] *private passenger auto liability*: a kind of insurance product which provides a comprehensive coverage protect against damages to a private automobile								
64	reinsurer	reinsurer profitability	professional reinsurer	[1] insurer or	reinsurer		size of the	reinsurer	
			ceding reinsurer						
	[1] reinsurer is a party that accepts all or part of the liabilities associated with the original policy issued by a primary insurance company.								

BIOGRAPHICAL DATA

NAME:	Mr. Tanu Khamphairoh
DATE OF BIRTH:	June 13, 1984
PLACE OF BIRTH:	Bangkok, Thailand
GRADUATION:	Bachelor's Degree of Arts (Business English) with Third Class Honors, Assumption University since 2006 Bachelor's Degree of Arts (Spanish), Ramkhamhaeng University since 2011
PRESENT POSITION:	PT (Personal Tutor) Wall Street Institute Thailand