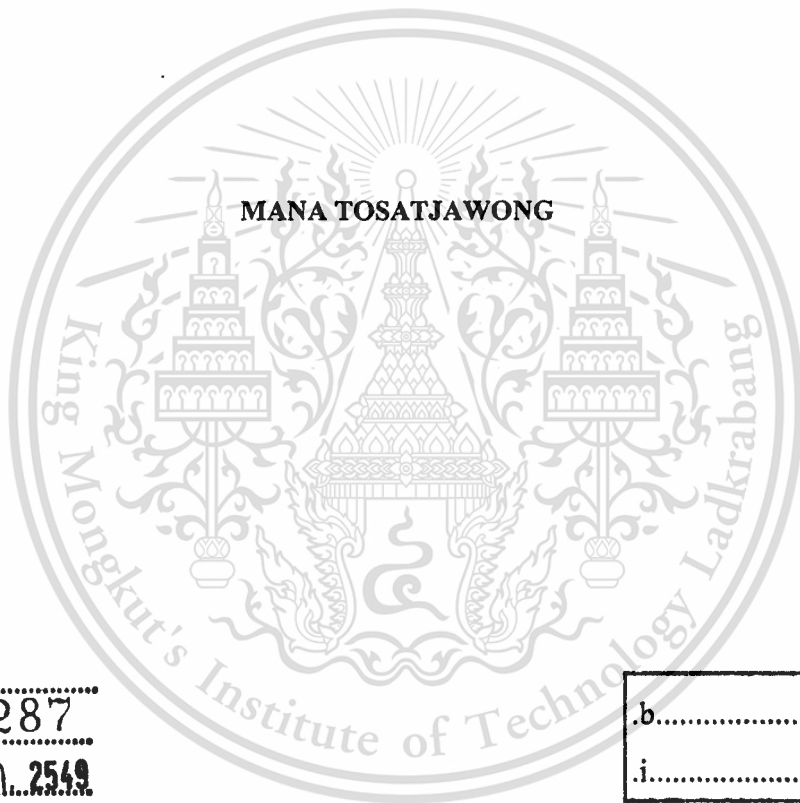


สำนักหอสมุดกลาง พระจอมเกล้าลาดกระบัง

**ANALYSIS OF VOCABULARY AND SENTENCE STRUCTURES
OF SPORTS NEWS IN THE BANGKOK POST**



เลขหมู่.....
เลขทะเบียน.....47287
วัน,เดือน,ปี...1.9...ค.ศ...2549

.b.....
.i.....

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

2006

ISBN 974-15-2753-5

This material is reserved for educational use only, not allowed for commercial use.

Forbidden to modify the content, and cite the document when use.

**ANALYSIS OF VOCABULARY AND SENTENCE STRUCTURES
OF SPORTS NEWS IN THE BANGKOK POST**



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

2006

ISBN 974-15-2753-5

This material is reserved for educational use only, not allowed for commercial use.

Forbidden to modify the content, and cite the document when use.



COPYRIGHT 2006

SCHOOL OF GRADUATE STUDIES

KING MONGKUT'S INSTITUTE OF TECHNOLOGY LADKRABANG

This material is reserved for educational use only, not allowed for commercial use.

Forbidden to modify the content, and cite the document when use.

หัวข้อวิทยานิพนธ์	การวิเคราะห์คำศัพท์และโครงสร้างประโยคของข่าวกีฬา ในหนังสือพิมพ์บางกอกโพสต์
นักศึกษา	นายมานะ โตสัจจะวงษ์
รหัสประจำตัว	44064364
ปริญญา	ศิลปศาสตรมหาบัณฑิต
สาขาวิชา	ภาษาศาสตร์ประยุกต์-ภาษาอังกฤษเพื่อวิทยาศาสตร์และเทคโนโลยี
พ.ศ.	2549
อาจารย์ผู้ควบคุมวิทยานิพนธ์	รศ. ดร. จิราภา วิทยากิริภัย
อาจารย์ผู้ควบคุมวิทยานิพนธ์ร่วม	ศศ. ดร. จิราภรณ์ อินทราไสย

บทคัดย่อ

วิทยานิพนธ์ฉบับนี้มีจุดประสงค์เพื่อวิเคราะห์คำศัพท์ และโครงสร้างประโยคที่ใช้ในข่าวกีฬาของหนังสือพิมพ์บางกอกโพสต์ ข้อมูลที่ใช้ในการวิเคราะห์คือข่าวกีฬาจำนวน 360 ข่าวซึ่งรวบรวมจากหนังสือพิมพ์บางกอกโพสต์เป็นเวลา 6 เดือนนับจากวันที่ 1 กรกฎาคม ถึง 31 ธันวาคม 2545 คลังข้อมูลข่าวกีฬาถูกแบ่งออกเป็น 7 กลุ่มย่อย เครื่องมือที่ใช้สำหรับกรวิเคราะห์คำศัพท์ได้แก่ซอฟต์แวร์คอนคอร์แดนซ์ชื่อ เวอร์ดสมิท รุ่นที่ 3 (WordSmith Tools Version 3) แนวทางการวิเคราะห์ประโยคอ้างอิงตามหนังสือ 'A University Grammar of English' ของเคิร์กและกรีนบาม (1973) และ 'Modern English: A Practical Reference Guide' ของมาร์เซลลา แฟรงค์ (1972)

ผลการวิเคราะห์คำศัพท์แสดงให้เห็นว่า คำเนื้อหาเกิดขึ้นจำนวนมากที่สุด (97.86 %) ตามมาด้วยคำย่อและชื่อย่อ (1.64 %) และคำไวยากรณ์ (0.50 %) ตามลำดับ คำว่า 'the' เกิดขึ้นมากที่สุด คำนามเฉพาะได้แก่ ชื่อนักกีฬาและประเทศพบมากในข่าวกีฬา คำกริยา 'said' ปรากฏในอันดับสูงที่สุด คำกริยาส่วนใหญ่เกิดขึ้นในรูปอดีตกาลมากกว่ารูปปัจจุบันกาล เนื่องจากการเล่าเหตุการณ์ข่าวกีฬาในอดีต คำบุพบทที่มีความถี่สูงที่สุดคือคำว่า 'in' คำสันธานที่เกิดขึ้นมากที่สุดคือ 'and'

สำหรับการวิเคราะห์โครงสร้างประโยค พบว่าจำนวนประโยคที่พบในข่าวกีฬามีทั้งหมด 3,612 ประโยค สังกรประโยค (Complex Sentences) เกิดขึ้นในความถี่สูงที่สุด 56.17 % ตามมาด้วยเอกัตถประโยค (Simple Sentences) 33.64 % อนกัตถ-สังกรประโยค (Compound-Complex Sentences) 6.81 % และอนกัตถประโยค (Compound Sentences) 3.38 % ตามลำดับ ผลการศึกษานี้สามารถนำไปเป็นแนวทางการศึกษาภาษาของข่าวกีฬาสำหรับนักเขียนข่าวกีฬา ผู้เขียนเนื้อหาและผู้ออกแบบหลักสูตร

Thesis Title	Analysis of Vocabulary and Sentence Structures of Sports News in the Bangkok Post
Student	Mr. Mana Tosatjawong
Student ID.	44064364
Degree	Master of Arts
Program	Applied Linguistics–English for Science and Technology
Year	2006
Thesis Advisor	Assoc.Prof. Dr. Jirapa Vitayapirak
Thesis Co-Advisor	Asst.Prof. Dr. Jiraporn Intrasai

ABSTRACT

This thesis aimed to analyze vocabulary and sentence structures used in the sports news in the Bangkok Post. The study used 360 sports news extracts from the Bangkok Post collected during six-month period of time from July 1st, 2002 to December 31st, 2002. The sports news corpus was divided into seven categories. Research tool for vocabulary analysis was concordancing software ‘WordSmith Tools Version 3’. The guidelines on the sentence structure analysis were based on ‘A University Grammar of English’ written by Quirk and Greenbaum (1973) and ‘Modern English: A Practical Reference Guide’ written by Marcella Frank (1972).

The results of vocabulary analysis showed that there were 97.86 % content words, 1.64 % abbreviations and acronyms, and 0.50 % function words, respectively. The word ‘the’ was the most frequent word in the sports news. Proper names, i.e. the names of sports players and countries, were frequently found in the sports news. The verb ‘said’ appeared at the highest rank among the other content words. Most verbs occurred in past tense because journalists used simple past to explain the past events in the sports news. Preposition with the highest frequency was ‘in’. The word ‘and’ was the most frequent conjunction.

Regarding the analysis of sentence structures, the total of sentence structures in the sports news was 3,612 sentences. Complex sentences occurred with the highest frequency (56.17 %), followed by simple sentences (33.64 %), compound-complex sentences (6.81 %), and compound sentences (3.38 %). The results of this study can be a guideline on the language of sports news for sports journalists, material writers, and course designers.

ACKNOWLEDGEMENT

First of all, I would like to express my sincere gratitude and deep appreciation to Assoc. Prof. Dr. Jirapa Vitayapirak, my respected advisor, for her invaluable guidance, useful and insightful comments and suggestions for the thesis as well as her supervision, patience and effort in reading my thesis. I am also grateful to Asst. Prof. Dr. Jiraporn Intrasai, my respected co-advisor, for her helpful advice, valuable time and supervision of the thesis.

I would like to thank my beloved parents, Mr. Komon and Mrs. Mina Tosatjawong, for their love and care, including my sister, Mrs. Anong Narong and my beloved friend, Ms. Siriporn Butnampetch, who comforted me when I felt tired from studying.

Next, I would like to thank all of my friends at King Mongkut's Institute of Technology Ladkrabang for their help, support, and encouragement.

Finally, I wish to thank those who have taken part in my success. They encouraged me to draw upon my inner strength to realize my dream of completing this thesis. Without them, my thesis would probably not have become a reality.

I will remember the time I have spent studying here forever.

Mana Tosatjawong

TABLE OF CONTENTS

	Page
THAI ABSTRACT	I
ENGLISH ABSTRACT	II
ACKNOWLEDGEMENT	III
TABLE OF CONTENTS	IV
LIST OF TABLES	VIII
LIST OF FIGURES	IX
CHAPTER 1 INTRODUCTION	1
1.1 Statement and Significance of the Problems.....	1
1.2 Goals and Objectives.....	2
1.3 Scope or Limitations of the Study.....	3
1.4 Process of the Study.....	3
1.5 Definitions.....	5
CHAPTER 2 LITERATURE REVIEW	7
2.1 Background of Sports News.....	7
2.1.1 News Sections of the Bangkok Post.....	7
2.1.2 Classification of Sports News.....	8
2.1.2.1 Advance Stories.....	8
2.1.2.2 Following Event Stories.....	8
2.1.2.3 Scores and Standings.....	8
2.1.2.4 Sports Activity Schedules.....	8
2.1.2.5 Journalist Comments.....	9
2.1.2.6 Reader Opinions.....	9
2.1.2.7 Interviewing Articles.....	9
2.2 Previous Research on the Language of Sports News.....	9
2.3 Vocabulary and Sentence Structures.....	10

This material is reserved for educational use only, not allowed for commercial use.

Forbidden to modify the content, and cite the document when use.

TABLE OF CONTENTS (CONTINUED)

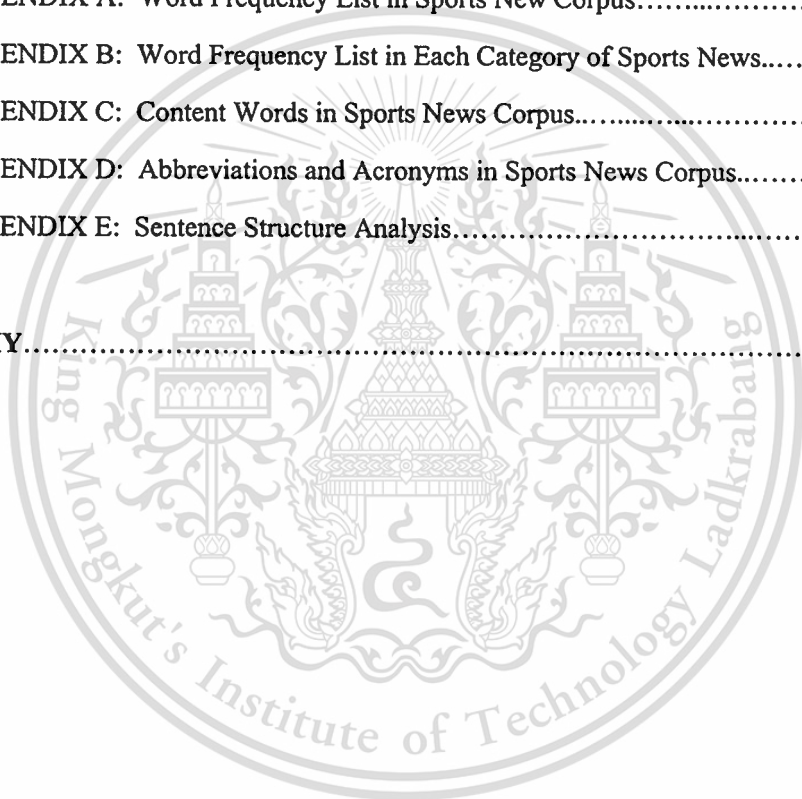
	Page
2.3.1 Vocabulary.....	10
2.3.1.1 Function Words.....	11
2.3.1.1.1 Prepositions.....	11
2.3.1.1.2 Conjunctions.....	11
2.3.1.2 Content Words.....	11
2.3.1.2.1 Nouns.....	11
2.3.1.2.2 Verbs.....	12
2.3.1.2.3 Adjectives.....	13
2.3.1.2.4 Adverbs.....	13
2.3.1.3 Abbreviations and Acronyms.....	14
2.3.2 Sentence Structures.....	14
2.3.2.1 Simple Sentences.....	15
2.3.2.2 Compound Sentences.....	19
2.3.2.3 Complex Sentences.....	20
2.3.2.4 Compound-Complex Sentences.....	23
2.4 A Corpus Study and Concordancing Software.....	24
2.4.1 A Corpus Study.....	24
2.4.1.1 Definition of Corpus.....	24
2.4.1.2 Types of Corpus.....	24
2.4.1.3 Corpus Design Considerations.....	25
2.4.2 Concordancing Software.....	26
2.4.2.1 Definition of Concordancing Software.....	27
2.4.2.2 Concordancing Software.....	27
2.4.2.3 Text Processing.....	27
CHAPTER 3 RESEARCH METHODOLOGY.....	31
3.1 Description of the Bangkok Post.....	31

TABLE OF CONTENTS (CONTINUED)

	Page
3.2 Process of the Study.....	31
3.2.1 Data Collection.....	31
3.2.2 Content Analysis.....	32
3.2.2.1 Vocabulary Analysis.....	32
3.2.2.2 Sentence Structure Analysis.....	32
3.2.3 Results and Discussions.....	33
3.2.4 Conclusions and Suggestions.....	33
CHAPTER 4 RESULTS AND DISCUSSIONS.....	34
4.1 Statistical Analysis of Sports News Corpus.....	34
4.2 Vocabulary Analysis.....	35
4.2.1 Function Words.....	35
4.2.2 Content Words.....	38
4.2.2.1 Top Twenty-Eight Content Words.....	38
4.2.2.2 Nouns, Verbs, Adjectives, Adverbs.....	40
4.2.3 Abbreviations and Acronyms.....	43
4.3 Sentence Structures Analysis.....	47
4.3.1 Simple Sentences.....	48
4.3.2 Compound Sentences.....	49
4.3.3 Complex Sentences.....	49
4.3.4 Compound-Complex Sentences.....	50
CHAPTER 5 CONCLUSIONS AND SUGGESTIONS.....	52
5.1 Conclusions.....	52
5.2 Implications for Journalists Education.....	53
5.2.1 Implications for Sports Journalists.....	53
5.2.2 Implications for Material Writers and Course Designers.....	53

TABLE OF CONTENTS (CONTINUED)

	Page
5.3 Suggestions for Further Studies.....	54
BIBLIOGRAPHY	55
APPENDICES	58
APPENDIX A: Word Frequency List in Sports New Corpus.....	59
APPENDIX B: Word Frequency List in Each Category of Sports News.....	79
APPENDIX C: Content Words in Sports News Corpus.....	83
APPENDIX D: Abbreviations and Acronyms in Sports News Corpus.....	88
APPENDIX E: Sentence Structure Analysis.....	92
BIOGRAPHY	107



LIST OF TABLES

Table		Page
1.1	Number of Sports News Stories Drawn for the Content Analysis.....	4
2.1	Statistical Details of Each Sub-field in Psychology Corpus.....	28
2.2	Word Frequency List of the Whole Corpus in the BNC.....	28
4.1	Statistical Details of Sports News Corpus.....	34
4.2	Top 10 Words in Sports News Corpus Compared with the Other Five Corpora.....	35
4.3	Top 14 Prepositions.....	36
4.4	Top 14 conjunctions.....	36
4.5	Top Twenty-Eight Content Words in Sports News Corpus Compared with the BNC...	38
4.6	Content Words in Sports News Corpus.....	40
4.7	Top 30 Words in Each Category of Sports News Corpus.....	41
4.8	Top 10 Abbreviations in Sports News Corpus.....	44
4.9	Examples of Abbreviations and Acronyms in Sports News Corpus.....	45
4.10	Number of Vocabulary in Sports News Corpus.....	46
4.11	Frequency of Occurrence of Sentence Structures.....	47

LIST OF FIGURES

Figure	Page
2.1 Concordance of ' <i>Factor</i> '.....	29
4.1 Concordance of ' <i>and</i> '.....	37



This material is reserved for educational use only, not allowed for commercial use.

Forbidden to modify the content, and cite the document when use.

CHAPTER 1

INTRODUCTION

1.1 Statement and Significance of the Problems

Sports news has been popular in English daily newspapers in great depth of coverage because sports are physical activities pursued for exercise or pleasure. In 1997, the number of pages for sports news in the Bangkok Post was increased from three pages to four pages. These pages report on competitions, the up-to-date news on teams or players they are interested in (Boonlikhit, 1998: 4-5). The reasons for the increased number of pages are that the readers wish to be freed or alleviated from various stresses, and the government and the private sector have paid an increased attention to sports business, especially soccer (Na Bangchang, 1996: 4-5).

There have been some studies of sports news in Thailand, but they were the studies of Thai daily newspapers, i.e. *The Analysis of Sports News Coverage in Daily Newspaper* (Boonlikhit, 1998), *Factors Related to the Growth of Sports* (Na Bangchang, 1996), and *An Analytical Study of Soccer News Translation in the Star's Soccer Daily Newspaper* (Sittirak, 2002). However, there has not been any research using the concordancing software to analyze the language features in the sports news corpus from English daily newspapers.

During the first half of the 19th century, the early corpora were produced to help language teachers trying to teach English systematically, and to simplify English for their learners (Jeffery, 1953: V). Studies of word frequency have become more interesting and challenging to English teachers and learners, with the development of the computer and information technology. Leech (1997: 10-11) described the advantages of using the computer to set up learning tasks in English courses. Firstly, the computer provides automatic searching, sorting, and scoring. Secondly, it can enable an individual's learning process. Finally, it encourages an exploration or discovery approach to learning.

In Thailand, there were very few studies oriented specifically to the systematic use of corpus-based information. In 1976, Boonya-udomsart conducted a comparative study between words in the General Service List and an Economics textbook. The result showed the number of new words that the students who studied Economics needed to know to understand the texts. Mallikamas (1999) exploited corpus analysis to create teaching materials and teaching pedagogy. She concluded that the use of corpora provides crucial and abundant data on linguistic description

This material is reserved for educational use only, not allowed for commercial use.

Forbidden to modify the content, and cite the document when use.

and valuable insights, which helped to improve students' knowledge and proficiency in English learning. Kaewphanngam (2002) studied "*A Corpus Analysis of Psychology Texts as a Basis for the Development of Teaching Materials in English for Academic Purposes*". He used the concordancing software to analyze the vocabulary in a corpus of psychology texts. The results showed that 'high frequency vocabulary', e.g. 'the', 'for', and 'in' was the largest group of vocabulary, whereas 'technical vocabulary' was the smallest group of vocabulary. Poonpon (2002) studied "*Vocabulary Input in English for Science Courses: A Corpus Analysis of Intensive and Extensive Course Materials*". She analyzed the vocabulary of course materials in English courses for first year and second year science students at Mahidol University and Khon Kaen University. This study revealed that vocabulary input in both intensive and extensive course materials in English for Science courses at Mahidol University and Khon Kaen University emphasized 'high frequency words'.

Given the problems already mentioned, this study attempts to compile a corpus of sports news text from the Bangkok Post which is the oldest and most popular English newspaper in Thailand (Phanphanich, 1999: 19). The sports news corpus is a specialized corpus. It is used to investigate a particular type of language in terms of vocabulary and sentence structures. The results of this study can be the source of vocabulary and sentence structures for mass communication students or persons who are interested in reading sports news from English newspapers.

1.2 Goals and Objectives

The purpose of the study is to analyze vocabulary and sentence structures used in sports news from the Bangkok Post. Three main objectives are as follows:

1.2.1 To analyze vocabulary based on sports news corpus.

1.2.2 To classify vocabulary into function words, content words, and abbreviations and acronyms.

1.2.3 To identify four types of sentence structures: simple, compound, complex, and compound-complex sentences.

1.3 Scope or Limitations of the Study

There are four limitations in this study.

1.3.1 The study analyzed vocabulary and sentence structures. The analysis of the vocabulary was done by using concordancing software '*WordSmith Tools Version 3.*' The analysis of the sentence structures was focused on simple, compound, complex, and compound-complex sentences and was based on '*A University Grammar of English*' written by Quirk and Greenbaum (1973) and '*Modern English: A Practical Reference Guide*' written by Marcella Frank (1972).

1.3.2 The sports news stories from the Bangkok Post were daily collected during a six-month period of time from July 1, 2002 to December 31, 2002.

1.3.3 The total sports news stories for the content analysis were 360.

1.3.4 Some sports news stories contain overlapping contents of more than one category. That is, the contents may be related to both an advance story and a following event story. The criteria for grouping such a sports news story depended on which category the contents are mostly based on.

1.4 Process of the Study

The process of the study was divided into four main stages, namely data collection, content analysis, results and discussions, and conclusions and suggestions.

1.4.1 Data Collection

The sports news stories of 5,375 were daily collected from the Bangkok Post during a six-month period of time. They were grouped into seven categories, namely advance stories, following event stories, sports activity schedules, scores and standings, journalist comments, reader opinions, and interviewing articles. The division of the categories is based on the study of Boonlikhit (1998) and an interview with Crutchley, the sports news editor of the Bangkok Post. The sports news stories in each category were calculated into percentages. From the percentage of the collection in each category, two sports news stories per day were randomly selected within the six-month period of time (180 days). Thus, the 360 selected sports news stories were drawn for the content analysis, as shown in Table 1.1.

Table 1.1 Number of Sports News Stories Drawn for the Content Analysis

Categories of Sports News	Sports News Stories Collected during 180 Days	Percentage	Sports News Stories for the Analysis
Advance Stories	1,623	30	108
Following Event Stories	3,166	58	209
Scores and Standings	219	4	14
Sports Activity Schedules	218	4	14
Journalist Comments	83	2	7
Reader Opinions	33	1	4
Interviewing Articles	33	1	4
Total	5,375	100	360

The 360 printed sports news stories were scanned and stored as plain-text files (*.txt) in the corpus by using Optical Character Recognition (OCR). As a preliminary check on a corpus size, there are 127,134 tokens or running words in the 360 sports news stories which are large enough to be the data for the analysis of a specialized corpus (Tribble and Jones, 1990).

1.4.2 Content Analysis

The content analysis was divided into two parts, i.e. vocabulary analysis and sentence structure analysis.

1.4.2.1 Vocabulary Analysis

The vocabulary was analyzed by using the concordancing software '*WordSmith Tools Version 3.*' The frequency of occurrence of the running words and word types, and type/token ratio were calculated. In this study, vocabulary analysis was divided into three types, i.e. function words (prepositions and conjunctions), content words (nouns, verbs, adjectives, and adverbs), and abbreviations and acronyms.

1.4.2.2 Sentence Structure Analysis

The sentences were identified and grouped into four main types, i.e. simple, compound, complex, and compound-complex sentences. The guidelines on the sentence structure analysis were based on '*A University Grammar of English*' written by Quirk and Greenbaum (1973) and '*Modern English: A Practical Reference Guide*' written by Marcella Frank (1972).

This material is reserved for educational use only, not allowed for commercial use.

Forbidden to modify the content, and cite the document when use.

1.4.3 Results and Discussions

The findings were shown and discussed according to the objectives of this study. The results of vocabulary analysis were presented in the form of frequency and percentage. The occurrence frequencies of sentence structures were calculated and the examples of the sentences were given. The results were then followed by discussions.

1.4.4 Conclusions and Suggestions

The conclusions and suggestions were presented in Chapter 5.

1.5 Definitions

BNC Corpus	BNC stands for British National Corpus. This 100-million-word corpus contains spoken (10 million words) and written (90 million words) British English that was created by Geoffrey Leech in 1995.
Brown Corpus	The one-million-word corpus of edited written American English that was produced by W. Nelson Francis and Henry Kucera in 1961.
Birmingham Corpus	The 20-million-word corpus of written English. This corpus was made by John Sinclair at the University of Birmingham in 1980.
Content Words	Group of words that reflected their story in the text. They include nouns, verbs, adjectives, and adverbs (Jurafsky and Martin, 2002: 314).
Function Words	Group of words that are used most extensively in the English language since they indicate the functions of language. They include prepositions and conjunctions (Jurafsky and Martin, 2002: 314).
Hapax Legomena	Words which occur only once in the corpus (singular hapax legomenon) (Kennedy, 1998: 100).
KWIC	This acronym stands for Key Word In Context. The KWIC appears in the center of each line with extra space on either side of it (Sinclair, 1991: 33, 173).
LLC	This acronym stands for London-Lund Corpus. The 500,000-word corpus of spoken British English from various genres has been prosodically transcribed. This corpus was initiated by Jan Svartvik in 1975.

This material is reserved for educational use only, not allowed for commercial use.

Forbidden to modify the content, and cite the document when use.

LOB Corpus	This acronym stands for Lancaster-Oslo-Bergen which is the corpus of one million words of written British English published in 1961 by Johansson et al.
Tokens	The number of all words encountered in the text. For example, the following sentence contains 15 tokens: ' <i>Michael Owen secured the winning goal for Liverpool in the 27th minutes at Vitesse Arnhem.</i> '
Sports News Stories	Sports news stories are found in the sports news section. The stories are divided into seven categories: advance stories, following event stories, sports activity schedules, scores and standings, journalist comments, reader opinions, and interviewing articles.
Vocabulary	A set of lexical items (words) in a language (Crystal, 1992: 412).
Word-form	The base plus all inflected, reduced, and derived forms (Sinclair, 1991: 41, 176). Hence <i>shoot</i> , <i>shoots</i> , <i>shooting</i> , <i>shot</i> , and <i>shooter</i> are different word-forms.
Word List	The list of vocabulary items in the corpus. The information can be shown in either alphabetical order or frequency order (Sinclair, 1991: 31).
Word Type	A token which occurred more than once in different positions was counted as one word type. For example, the sentence ' <i>Michael Owen secured the winning goal for Liverpool in the 27th minutes at Vitesse Arnhem</i> ' contains 15 tokens, but 14 word types since <i>the</i> occurs twice.

CHAPTER 2

LITERATURE REVIEW

This chapter draws on the background of sports news, an overall theoretical framework of vocabulary and sentence structures, and corpus study. The chapter begins with the background of sports news. The second section reviews the previous research in sports news. The third part presents vocabulary and sentence structures. The final section describes corpus study and concordancing software.

2.1 Background of Sport News

The increase of the public interest on sports has affected newspapers. People pay a lot of attention to sports events. Some readers turn to sports sections before reading important front page news. Jaumsupanimit (1986: 183) states that sports editors always make sure that their publications cover sports and attract the largest audience or most of their readers. The readers' purposes in reading sports news are as follows:

- To follow news about the sports or players they are interested in.
- To find out how to improve their own games.
- To learn about neighbourhood teams they can join.
- To find sports events they can join.

2.1.1 News Sections of the Bangkok Post

It is essential to know the overall sections in the Bangkok Post. Kraisornkovit (1997: 99-108) classifies the news coverages in the Bangkok Post into four main sections as follows:

- Section I - News

This section consists of home news, regional news, international news, opinion & analysis news, and sports news.

- Section II - Business

This section consists of home business news, international business news, stock exchange news, investment news, and international stock markets news.

- **Section III - Outlook and Real time**

The Outlook consists of daily column, i.e. society, features, comics & puzzles, briefings, and television. Furthermore, this section also provides the column in a week, i.e. art, entertainment, health & fitness, nature, people, books, and food. The Real Time consists of music, society, comics & puzzles, columns, movies, dine & wine, what's on, TV highlight, television, at home, and features.

- **Section IV - Separated News**

This section consists of scorecard, learning post, database, horizons, motoring, perspective and Sunday comics.

In this study, the sports news stories are taken from Section I of the Bangkok Post which consists of four to five pages.

2.1.2 Classification of Sports News

According to Boonlikhit (1998: 10-11) and an interview with Roger Crutchley (August 18, 2002), sports editor of the Bangkok Post, the sports news are classified into seven categories as follows:

2.1.2.1 Advance Stories

The advance stories refer to the important competition, the criticism of playing, the comparison of competition in the past, the playing system of each team, and the popular players.

2.1.2.2 Following Event Stories

The following event stories involve the competition results, the comparison of each team, and reports of players or teams.

2.1.2.3 Scores and Standings

The scores and standings concern points, goals, or scores made by a player or team in a competition.

2.1.2.4 Sports Activity Schedules

The sport activity schedules explain plans or lists of events together with the times.

2.1.2.5 Journalist Comments

The journalist comments provide opinions about or explanations for the competitions and players depending on facts and reliable evidence.

2.1.2.6 Reader Opinions

The reader opinions allow audience for asking various problems and expressing their opinions.

2.1.2.7 Interviewing Articles

The interviewing articles provide conversations in which well-known persons such as players, team managers, coaches, team controllers, committees talk to a reporter about the interesting or important matters of sports.

The contents of the sports news above reflect the range of sports news in the Bangkok Post. This classification is used as the basis for deciding what sports areas to include in sports news corpus. In other words, the contents of the sports news corpus are designed to reflect the range of news sections of the Bangkok Post.

2.2 Previous Research on the Language of Sports News

Research on sports news stretch from the pioneer works to recent investigations. There has been a substantial amount of research on the study of the language used in the news. Some research studied news in English daily newspaper such as Front Page News (Thanomsak, 1998), Political News (Phanphanich, 1999), and Technology News (Tang-U-Thaisuk, 2000). Some research were the studies of sports news in Thai daily newspapers, i.e. The Analysis of Sports News Coverage in Daily Newspaper (Boonlikhit, 1998), Factors Related to the Growth of Sports (Na Bangchang, 1996), and An Analytical Study of Soccer News Translation in the Star's Soccer Daily Newspaper (Sittirak, 2002). The following section reviews significant research related to the English language used in the news.

Phanphanich (1999) studied the general characteristics of the language used in political news and analyzed pieces of political news in English daily newspapers. The result was classified into four categories, namely sentence complexity, omission of the verb to be in present and past participles, use of the active and passive voice, and register. The findings of the language analysis showed that the complex sentence occurred at the highest rank (43.40%), followed by simple

This material is reserved for educational use only, not allowed for commercial use.

Forbidden to modify the content, and cite the document when use.

sentence (30.65%), compound sentence (16.42%) and compound-complex sentence (9.53%). The frequency of occurrence of active voice was greater (88.73%) than the passive voice (11.27%). The registers were used to clarify the vocabulary which occur and convey meaning in the political news.

There were, in addition, some studies abroad on grammatical structures and vocabulary of news. Timothy N. Walters (1994) studied "*Syntax and Successful Placement of Press Releases in Newspapers.*" He observed the grammatical structures of original press releases in daily newspapers. He summarized that most of the journalists were aware of grammatical structures and made them simpler, shorter, easier to read, and less passive. Thompson (1991) studied about "*A Content Analysis of Newspaper Use of Language*". He found that specific vocabularies were constructed for 'oral' words (including said, announced, discussion), 'print' words (including ballot, law, note, wrote), and 'education' words (including college, diploma, education, and professor). The results indicated that 'oral' and 'education' words were used more frequent than 'print' words. Furthermore, Monda et al. (1988) quoted in Aiex (1988) that newspaper could be helpful for the teachers who wanted to create an individual program by using specific learning skills in reading. They also suggested that teacher should use words in grocery ads, sports section, or comics to teach alphabetizing, and use news stories to teach grammar.

It is clear from the above findings that almost all of the previous research did not concern the sports news in English newspapers. Thus, it is worthwhile carrying out this research to find out the vocabulary and sentence structures in sports news.

2.3 Vocabulary and Sentence Structures

This section is divided into two parts: vocabulary and sentence structures.

2.3.1 Vocabulary

Vocabulary is a set of lexical items in a language (Crystal, 1992: 412). They are classified according to their functions which are used to form sentences (Frank, 1972: 1). In this study, vocabulary analysis was divided into three types, i.e. function words, content words, and abbreviations and acronyms.

2.3.1.1 Function Words

Function words are the group of words that are used most extensively in the English language since they indicate the functions of language. They include prepositions and conjunctions (Jurafsky and Martin, 2002: 314).

2.3.1.1.1 Prepositions

A preposition is a word such as *by, for, into, or with*, which usually has a noun group as its object. Prepositions include one point of time (*on, at, in*); extended time (*since, by, from-to, until, for, during, in or within*); sequence of time (*before, after*); the point itself (*in or inside, on, at*); higher or lower than a point (*over, above, under, underneath, beneath, below*); neighboring the point (*near, next to, alongside, beside, between, opposite*).

2.3.1.1.2 Conjunctions

A conjunction is the member of a small class that has no characteristic form. It functions as nonmovable structure word that joins such unit as parts of speech, phrases, or clauses (Frank, 1972: 206-08, 215). There are two types of conjunction, i.e. coordinate and subordinate conjunction. A coordinate conjunction joins structural units that are equal grammatically, e.g. *and, or, but, yet*. A subordinate conjunction introduces a clause that depends on a main, or independent clause, e.g. *when, until, if, that, whether, before, after*.

2.3.1.2 Content Words

Content words belong to open classes in the sense that it is indefinitely extendable. They include nouns, verbs, adjectives, and adverbs (Jurafsky and Martin, 2002: 314).

2.3.1.2.1 Nouns

A noun is one of the most important parts of speech. Its arrangement with the verb helps to form the sentence core which is essential to every complete sentence (Frank, 1972: 6-7). Nouns are classified according to meaning, i.e. proper nouns, common nouns, concrete nouns, abstract nouns, countable nouns, noncountable nouns, and collective nouns, or form, i.e. noun compounds.

A proper noun begins with a capital letter in writing. It includes personal names (*John Smith, Susan Brown*), name of geographic unit such as countries, cities, rivers, etc. (*England, New York, Niles*), names of nationalities and religions (*Dutchman, Christianity*), names of holidays (*Easter, Thanksgiving Day*), names of time units (*Saturday, January*), and words used for personification—a thing or abstraction treated as a person (*Nature, Liberty*). As opposed to proper

nouns, all other nouns are classified as common nouns (book, table, window). A concrete noun is a word for a physical object that can be perceived by the senses (flower, girl, cat). Abstract noun is a word for a concept. It is an idea that exists in our minds only (truth, freedom, guilt). Countable noun can usually be made plural by the addition of -s (one car, two boys, three pens). Noncountable noun is not used in the plural (water, weather, heat). Collective noun is a word for a group of people, animals or subjects considered as a single unit (crowd, class, committee, group, team). Noun compounds refer to a group of words joined together into one vocabulary unit that functions as a single part of speech (bathroom, lady's maid, blue print, dance team, lifeguard).

2.3.1.2.2 Verbs

A verb is a word or a phrase which express the existence of a state or the doing of an action. Verbs can be classified by complement of verbs, i.e. predicating or linking verbs, transitive or intransitive verbs and reflexive verb, and by form of verbs, i.e. auxiliary or lexical verbs, and finite or non-finite verbs.

A predicating word (or predicator) has been called a verb of action (They *shout*; He *signed* the contract.), but has been interpreted to include non-action verbs that are not linking verbs (I *remember* him; They *need* more money). The common linking verbs are *appear, be, become, get, look, remain, seem*.

A transitive verb takes a direct object (He *is reading* a book). An intransitive verb does not require an object (He *is walking* in the park). Only transitive verbs may be used in the passive voice (The book was returned by him quickly). Many verbs may be used either transitively (He *was writing* a letter) or intransitively (He *was writing* at the library). Other verbs are transitive only and must be followed by an object, e.g. *have, like, need, owe, remember*.

Reflexive verb requires one of the compounds with -self (reflexive pronoun) as its object (express oneself, pride oneself). Some verbs may be used with or without the reflexive pronoun object, e.g. He washed (himself) and dressed (himself) quickly.

Auxiliary (or auxiliaries) is the first part of the verb phrase (*is opening*) and the second part is the lexical verb (*is opening*). Some verbs used as auxiliaries may occur independently (be, have, do). Other verbs function as auxiliaries and must be used with a lexical verb (shall, will, can, must, may, should, would, could, might). Auxiliaries combine with lexical verbs that are in simple (unchanged) form (will *score*), in -ing form (*is scoring*), or in -ed form (*was scored*).

A finite verb is a lexical verb with or without auxiliaries that functions as the full verb in the predicate (He tries to keep the ball). Non-finite (or infinite) verbs are incomplete verb forms that function as other parts of speech than verbs. They consist of the participle -ing or -ed forms and infinitive forms -to + the simple form of the verb (The boy *talking to the teacher* is my brother; He tries *to keep the ball*).

2.3.1.2.3 Adjectives

The adjective is a modifier that has the grammatical property of comparison. It is identified by special derivational endings or by special adverbial modifiers that precede it (Frank, 1972: 109-11). Adjective is classified into two types, i.e. determiners and descriptive adjectives.

Determiners consist of a small group of structure words without characteristic form. They include articles (the, a, an); demonstrative adjectives (this, these, that, those); possessive adjectives (my, your, one's, the player's); numeral adjectives (four, twenty-five, fourth, twenty-fifth); adjectives of indefinite quantity (some, few, all, more); relative and interrogative adjectives (whose, what, which).

Descriptive adjectives usually indicate an inherent quality or a physical state, such as age, size, color. Inflectional and derivational endings can be added only to this type of adjective. Some descriptive adjectives take the form of proper adjectives (an Australian swimmer, an Evertonian play); participle adjectives (an interesting game); adjective compounds (a good-looking fan, a broken-down stadium); adjective with -ed added to nouns functioning as the second element of a compound (far-sighted, ill-tempered); adjective compounds that denote parts of the body (left-handed, kind-hearted, bare-footed); adjective compounds that include various syntactic combinations (an up-to-the-minute office, hard-to-please employer, a life-and-death struggle); some phrases that function as adjective compounds (a catch-as-catch-can policy); adjective compounds that express measurement (a ten-foot-deep-hole, a sixteen-year-old boy, a hole ten feet deep, a boy sixteen years old); noun compounds that function as adjective compounds (a high-school competition, twentieth-century ceremony).

2.3.1.2.4 Adverbs

Adverbs range in meaning from words having a strong lexical content to those that are used for emphasis. They range in function from close to loose modifiers of the verb (Frank, 1972: 141-45). Adverbs can be classified by meaning, i.e. manner, place and direction, time, and intensifying adverbs.

The manner adverb has the most characteristic adverbial form, an –ly ending added to a descriptive adjective (I began to run quickly). The adverbs of place and direction may include some prepositional forms appearing after the verb (He went outside; They walked down.). Definite time adverbs have a fixed boundary in time (yesterday, today, tomorrow). Indefinite time adverbs have no fixed boundary (recently, nowadays, soon, now, then, before, after, always, often, sometimes). Intensifying adverbs are quantifiers (very, too, rather, quite, fairly, more) or emphasers (almost, entirely, nearly, partially).

Adverbs can be classified by function, i.e. sentence adverbs (fortunately, actually, obviously, evidently); conjunctive adverbs (therefore, moreover, however, otherwise, then); explanatory adverbs (namely, for example, as, i.e., e.g., viz); relative and interrogative adverbs (when, where, why, how); exclamatory adverbs (how).

2.3.1.3 Abbreviations and Acronyms

An abbreviation is any shortened form of a word or phrase used to present the complete form, e.g. *Tues.* for *Tuesday*; *Rd* for *Road*; *exam* for *examination*. A contraction is a kind of abbreviation. It may be shortened forms of phrases which include the first letter of the first word and the last letter of the last word, e.g. *I'll* for *I will*, *isn't* for *is not*, *it's* for *it is*. An acronym is a specific kind of abbreviation. It can be either a pronounceable word formed from the initial letters of a name, e.g. *NATO* for *North Atlantic Treaty Organization*; *OPEC* for *Organization of Petroleum Exporting Countries*, or a combination of initial letters or parts of a series of words, e.g. *radar* for *radio detection and ranging* (Ajjan, 1993: x).

2.3.2 Sentence Structures

In this study, the guidelines on the sentence structure analysis were based on '*A University Grammar of English*' written by Quirk and Greenbaum (1973) and '*Modern English: A Practical Reference Guide*' written by Marcella Frank (1972). The classifications of sentences are based on the number and kinds of clauses. A clause may be defined in the same way as a sentence. It is a full predication that contains a subject and a predicate with a finite verb. There are two kinds of clauses, i.e. independent and dependent clauses. An independent clause is a full predication that may stand alone as a sentence. A dependent clause has a special introductory word that makes the predication depending on an independent clause. Frank (1972: 222-223) divides sentence structures into four main types: simple, compound, complex and compound-complex sentences.

2.3.2.1 Simple Sentences

A simple sentence has only one full predication in the form of an independent clause (Frank, 1972: 222). It has one finite verb. As provided by Quirk and Greenbaum (1973: 167), simple sentences can be divided into seven basic clause types as follows:

- | | |
|----------|--|
| (1) SV | Subject + Verb
Someone is talking. |
| (2) SVA | Subject + Verb + Adverbial
My car is outside the house. |
| (3) SVC | Subject + Verb + Subject Complement
Susan is my accountant.
My brother is ill. |
| (4) SVO | Subject + Verb + Object
The man stole the jewelry. |
| (5) SVOO | Subject + Verb + Indirect Object + Direct Object
She gave me the letter. |
| (6) SVOA | Subject + Verb + Object + Adverbial
You can put your coat in my bedroom. |
| (7) SVOC | Subject + Verb + Object + Object Complement
You have made me very happy. |

In case of the sentences consisting of compound subjects as shown in (8), compound predicates in (9), and compound objects in (10), such sentences were analyzed as one independent clause or a simple sentence.

(8) Dogs and cats often do not get along.

(9) They went out for dinner but returned in time for their favorite television program program.

(10) He looked everywhere for his keys and wallet.

According to Quirk and Greenbaum (1973: 170-171), the five sentence elements that appear in the seven basic clause types are *subject*, *verb*, *object*, *complement*, and *adverbial*.

- Subject (S)

A subject is a noun phrase with nominal function. It occurs before the verb phrases in declarative clauses, and after the auxiliary verbs (be, have, do) or modal verbs (shall, will, can, must, may, should, would, could, might) in questions.

(11) Mark has written a good paper.

S

(12) Do you want to go shopping?

S

(13) Could I borrow your pencil?

S

- Verb (V)

A verb is a verb phrase that follows the subject and precedes the object, complement, or adverbial. A transitive verb (tv) takes a direct object whereas an intransitive verb (iv) does not require an object.

(14) He is reading a book.

tv

(15) He is walking in the park.

iv

With regard to passive clauses, a passive verb phrase is symbolized here as V_{pass} . The various clause types can be transformed into the passive clauses by means of transformational relations, or relations of grammatical paraphrase. Thus, there is equivalence between the clause types and their passive clauses, as illustrated below.

(16) SVO \leftrightarrow SV_{pass} (A)

Many critics disliked the play. \leftrightarrow The play was disliked (by many critics).

(17) SVOC \leftrightarrow $SV_{\text{pass}}C$ (A)

Queen Victory considered him a genius. \leftrightarrow He was considered a genius (by Queen Victory).

(18) SVOA ↔ SV_{pass}A (A)

An intruder must have placed the ladder there. ↔ The ladder must have been placed there (by an intruder).

(19) SVO_iO_d ↔ SV_{pass}O_d (A)

or SV_{pass}O_i (A)

My father gave me this watch. ↔ I was given this watch (by my father).

↔ This watch was given me (by my father).

Regarding phrasal verbs, an intransitive phrasal verb is a multi-word verb consisting of a verb plus a particle. The following is an example of an intransitive phrasal verb.

(20) The plane has now taken off.

S (A) V

Moreover, the phrasal verb can take a direct object. As seen from the examples (21) and (22), with most transitive phrasal verbs, the particle can either precede or follow the direct object.

(21) They turned on the light.

S V O

(22) They turned the light on.

S V O A

The meaning of the combination of a verb and a particle cannot be predicted from the meanings of the verb and particle in isolation, as illustrated in (23) and (24).

(23) His father gave up smoking two years ago.

S V O (A)

(give up = stop doing something)

(24) Richard had turned up on Christmas Eve with Tony.

S V (A) (A)

(turn up = arrive unexpectedly or after one have been waiting a long time)

Besides the seven basic clause types mentioned above, there-existential sentences derived from regular clauses by means of a general rule can be analyzed as a simple sentence. That is,

there is a regular relation of equivalence between the existential clauses and the basic clause types. The general rule (Quirk and Greenbaum, 1973: 418) is presented below:

Subject + (auxiliaries) + be + predication

→ There + (auxiliaries) + be + subject + predication

Examples of there-existential sentences are shown in (25), (26), (27), (28), (29), (30) and (31).

(25) SVC: Something must be wrong. → There must be something wrong.

(26) SVA: Was anyone around? → Was there anyone around?

(27) SV: No one was waiting. → There was no one waiting.

(28) SVO: Plenty of people are getting promotion. → There are plenty of people getting promotion.

(29) SVOC: Two bulldozers have been knocking the place flat. → There have been two bulldozers knocking the place flat.

(30) SVOA: A girl is putting the kettle on. → There's a girl putting the kettle on.

(31) SVOO: Something is causing her distress. → There's something causing her distress.

In addition, Quirk and Greenbaum (1973: 419) stated that the 'bare existential' sentence simply shows the existence of some entity or entities. It has a simple clause structure *there + be + indefinite noun phrase*, as shown in (32) and (33).

(32) Undoubtedly, there is a god. (= God exists)

(33) There have always been wars. (= Wars have always existed/ taken place.)

- Object (O)

An object, like a subject, is a noun phrase with nominal function. It follows the subject and the verb phrase. There are two types of object, i.e. a direct object (O_d) and an indirect object (O_i). An indirect object always precedes the direct object.

(34) Nancy showed me her book.

O_i O_d

- Complement (C)

A complement is a noun phrase or an adjective phrase, having a co-referential relation with a subject or an object. A subjective complement (C_s) comes after a linking verb to refer back to the subject and has the same identity as the subject. An objective complement (C_o) follows the direct object and has the same identity as direct object.

(35) He is the chairman.

C_s

(36) They make him the chairman.

C_o

Subject and object complements have the additional possibility of being realized by adjective phrases (having an adjective as head), as in (37) and (38).

(37) My brother is ill.

C_s

(38) She made him happy.
 { C_o
very much happier.

C_o

- Adverbial (A)

An adverbial can be realized by adverb phrases, by noun phrases, or by preposition phrases, as shown in (39), (40) and (41), respectively.

(39) John very carefully searched the room.

(40) They make him the chairman every year.

(41) The girl studied at a large university.

2.3.2.2 Compound Sentences

A compound sentence has two or more full predications in the form of independent clauses (Frank, 1972: 223). The full predications may be joined coordinately by a semicolon (;), by coordinate conjunctions (and, or, but, nor, yet, so), by conjunctive adverbs (then, therefore, hence, however), or by correlative conjunctions (either...or, neither...nor, both...and, not only...but also). The followings are examples of compound sentences.

This material is reserved for educational use only, not allowed for commercial use.

Forbidden to modify the content, and cite the document when use.

- (42) Tim was watching television; his sister was on the phone.
- (43) The man stole the jewelry and he hid it in his home.
- (44) John was sick; therefore, he didn't come to school.
- (45) She will either pay for the ring or she will return it.

2.3.2.3 Complex Sentences

A complex sentence has two or more full predications. One of these is an independent clause (or main clause) that is a simple sentence, and one or more of these are dependent clauses (or subordinate clauses) (Frank, 1972: 223). Dependent clauses are in the form of finite or non-finite clauses which subordinate to the main clauses. Finite clauses consist of *adverbial clauses*, *adjective clauses*, or *noun clauses*. Non-finite clauses consist of *-ing participle*, *-ed participle*, *to-infinitive*, *verbless clause* and *bare infinitive*.

An adverbial clause is a subordinate clause that functions as adverbial in sentence structure and consists of a subject and a predicate introduced by a subordinate conjunction like *when*, *while*, *after*, *although*, *because*, or *if*. A clause may be used in an initial position, a mid-position, or a final position. When adverbial clauses are punctuated, the punctuation marks are commas. In mid-position, the clauses are enclosed in a pair of commas. The followings are examples of adverbial clauses in complex sentences.

- (46) *While my father was walking home*, he saw an accident.

Dependent Clause *Independent Clause*
(*Adverbial Clause*)

- (47) My parents, *needing money for house repairs*, applied for a mortgage.

↑ *Dependent Clause (-ing participle)* ↑
_____ *Independent Clause* _____

- (48) The child stayed at home *because he was ill*.

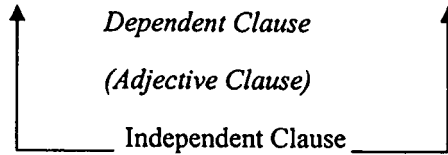
Independent *Dependent Clause*
Clause (*Adverbial Clause*)

An adjective clause is a clause that functions as a post-modifier in a noun phrase. In the adjective clause, a full subject and predicate is changed by means of a special introductory word like *who*, *whose*, *where*, *which* or *that*, which has the same referent as the preceding noun or pronoun.

This material is reserved for educational use only, not allowed for commercial use.

Forbidden to modify the content, and cite the document when use.

(49) The man *who stole the jewelry* hid it in his home.



(50) This is the girl *whose picture you saw*.

Independent *Dependent Clause*
 Clause (*Adjective Clause*)

(51) I am in the new house *where he lives*.

Independent *Dependent Clause*
 Clause (*Adjective Clause*)

A noun clause is a subordinate clause that functions as a noun and consists of a subject and a predicate. The following are the examples of noun clauses.

(52) I know *that coffee grows in Brazil*.

Independent *Dependent Clause*
 Clause (*Noun Clause*)

(53) Alex told us *how he will get the money?*

Independent *Dependent Clause*
 Clause (*Noun Clause*)

(54) He suggested *that I write the letter soon*.

Independent *Dependent Clause*
 Clause (*Noun Clause*)

(55) I hadn't realized *what a pretty girl she was*.

Independent *Dependent Clause*
 Clause (*Noun Clause*)

Non-finite clauses which are -ing participle, -ed participle, to-infinitive, verbless clause, and bare infinitive can be regarded as reduced clauses. -Ing participle, -ed participle, and to-infinitive lack the subjects and have non-finite verbs. A clause may be used in an initial position, a mid-position, or a final position. When non-finite clauses are punctuated, the punctuation marks are commas. In the mid-position, the clauses are enclosed in a pair of commas. The following are examples of non-finite clauses.

This material is reserved for educational use only, not allowed for commercial use.

Forbidden to modify the content, and cite the document when use.

Complex sentences, moreover, can be found in the forms of postponement and a cleft sentence.

- Postponement

A clausal subject is placed at the end of the sentence, and the subject position is filled by the anticipatory pronoun “*it*”. The resulting sentence contains two subjects, which are the postponed subject (the clause which is the subject of the sentence) and the anticipatory subject “*it*”. The followings are the examples of the postponement.

(65) It’s a pity *to make a fool of yourself*.

(cf: To make a fool of yourself is a pity.)

(66) It surprised me *to hear him say that*.

(cf: To hear him say that surprised me.)

- Cleft Sentences

In cleft sentences, the sentence is divided into two and one part is given greater prominence. The sentence begins with the pronoun “*it*” followed by the verb “*be*”, which in turn is followed by the element on which the focus falls. From a single clause such as “*John wore his best suit to the dance last night*”, it is possible to derive four cleft sentences, each highlighting a particular element of the clause, as shown in the following examples.

(67) S as focus: It was *John* who/that wore his best suit to the dance last night.

(68) DO as focus: It was *his best suit* (that) John wore to the dance last night.

(69) A_{time} as focus: It was *last night* (that) John wore his best suit to the dance.

(70) A_{place} as focus: It was *to the dance* that John wore his best suit last night.

2.3.2.4 Compound-Complex Sentences

A compound-complex sentence contains two or more independent clauses and one or more dependent clauses (Frank, 1972: 223). The following are examples of compound-complex sentences.

(71) [The man stole the jewelry] and [he hid it in his home] *until he could safely get*

Independent	Conj	Independent	Dependent Clause
Clause 1		Clause 2	(Adverbial Clause)

out of the town.

This material is reserved for educational use only, not allowed for commercial use.

Forbidden to modify the content, and cite the document when use.

(72) Bill said *he met Susan yesterday* and she gave him a present.

Independent	<i>Dependent Clause</i>	Conj	Independent
Clause 1	<i>(Noun Clause)</i>		Clause 2

2.4 A Corpus Study and Concordancing Software

This section presents the background and descriptions of the corpus study and concordancing software used for vocabulary analysis in this study.

2.4.1 Corpus Study

The background and descriptions of the corpus study can be divided into three parts: definition of corpus, types of corpus, and corpus design considerations.

2.4.1.1 Definition of Corpus

A corpus is a large and principle collection of natural text. A corpus (plural: corpora) is a collection of linguistic data, either written texts or spoken texts, and usually stored in a computer-readable form (Boswood, 1997: 253). It is the basis for analyzing the actual patterns of use in natural texts and go beyond the simple count of linguistic features (Biber et al., 1998: 4-5).

2.4.1.2 Types of Corpora

Corpora are designed according to their purposes, contents, and formats (Arunmanakul, 2002: 18-19). Regarding the purposes, Hunston (2002: 14) states that corpora are categorized into two ways, i.e. general and specialized corpus. Firstly, general corpus is a large corpus of texts of many types. It includes written or spoken language, or both (Sinclair, 1991: 17). Large corpora include the BNC, COBUILD, Brown, and LOB Corpus. Secondly, specialized corpus is a corpus of texts of a particular type, such as newspaper editorials, geography textbooks, academic articles, and so on. It is used to investigate particular types of language. A general corpus is usually much larger than a specialized corpus (Hunston, 2002: 15). In addition, Gallerstam (1991: 156) notes that corpora for general purposes are relatively large and corpora tailored for special purposes are relatively small.

Corpora are classified into written or spoken corpora according to their contents. Hartmann (2001: 179-180) suggests that written corpus is a collection of written texts which document the distinctive features of a language variety for use as data in lexicographic projects and dictionary

research. Spoken corpus is a collection of transcriptions of sound recording to document the usage of speech.

Arunmanakul (2002: 19) comments that there are basically two types of corpora, i.e. plain and annotated. Plain texts corpus consists of complete texts without annotation and can be prepared by using a word processor. It is to keep the text unprocessed and clean of any other codes (Sinclair, 1991: 26). Annotated texts corpus contains a minimum of annotation which identifies divisions as paragraphs and line number. The texts can be annotated to show the word class in the text by tagging to each word (Kennedy, 1998: 21).

In this study, sports news corpus was stored as plain texts. With respect to the purpose, sports news corpus was classified into specialized corpus because it is relevant to one particular interest or discipline (Shei, 1999: Online), i.e. sports news collected from English daily newspaper.

2.4.1.3 Corpus Design Considerations

Compiling the electronic text corpus, we need to design the criteria for major decisions and establish the framework. The corpus design considerations are objectives, getting permission, selection criteria for source materials, representativeness and balance, sampling, and corpus size.

The first consideration in creating corpus is the purposes of corpus design. The purposes of specialized corpora can be much more specific than general corpus. The other questions of decision include whether what is to be in the corpus, a corpus should be static or dynamic, and written or spoken corpus, how it can be representative and how big a corpus should be (Kennedy, 1998: 60).

Kennedy (1998: 76) suggests that corpus compilers must follow the same rules and legal requirement. Before texts are copied into a corpus database, compilers must seek and gain the permission of the authors and publishers who hold copyright for a work.

Corpus should present a static or dynamic picture of a language (Kennedy, 1998: 60). It can be a static collection of texts selected in some principled way and intended to be typical of the whole language or an aspect of the language at a particular time. Dynamic or monitor corpus (Sinclair, 1991; 25) provides the means to monitor changing patterns of usage over time. The vast amounts of text are collected over a period of time and passed through software. A dynamic collection of texts is growing and changing with the addition of new text samples (Kennedy, 1998: 61).

This material is reserved for educational use only, not allowed for commercial use.

Forbidden to modify the content, and cite the document when use.

A crucial difference between general and specialized corpus design is the method of sampling. A specialized corpus should involve the target of broad matter area by a non-random selection which concerned with the diversity of text that subcategories include. The 'proportional sampling' is one kind of non-random selection. It involves representing the different categories of texts (Biber et al., 1998: 246).

The issue of how many tokens of a linguistic item is necessary for descriptive adequacy. Sinclair (1991: 18) guides that a corpus should be as large as possible and keep on growing. In order to study the behavior of words in the texts, compilers need to have quite a large number of occurrences. Whereas, Biber et al. (1998: 269) comments that text samples of 2,000-5,000 words are big enough to represent their text categories. Kennedy (1998: 68) also argues that there is no point in having bigger and bigger corpora if you cannot work with the output. He emphasizes that a huge corpus does not necessarily represent a language or a variety of a language any better than a small corpus.

Rather than focusing on the quantity of data in a corpus, compilers and analysts need to know that the quality of the data they work with is important. A corpus is more or less adequate according to the extent to which the corpus matches the purposes to which it is put (Kennedy, 1998: 68). Moreover, a small corpus might be sufficient if it can produce the interesting results. In contrast, large corpus might not be sufficient if they fail to represent the range of variation (Biber et al., 1998: 169).

Carter and McCarthy (1995: 143) argue that for the purposes of studying grammar in spoken language, a small corpus is sufficient because grammar tend to be very frequent. For the study of prosody, a corpus of 100,000 words will usually be big enough to make generalizations for most descriptive purposes.

The descriptions above review the principal aspects of corpus creation and the major decisions to be made when creating an electronic text corpus.

2.4.2 Concordancing Software

The concordancing program is a kind of computer program which is used to investigate the language features of given texts. In recent year, the computer programs have become increasingly available. It is a set of instructions to a computer and can be very useful for analyzing corpus.

2.4.2.1 Definition of Concordancing Software

Concordancing software is the program which tells the computer how to analyze a corpus. It tells what texts use as input, what linguistic features to analyze and how to identify them, and what kind of output to produce (Landau, 2001: 254). In other words, the concordancing program is the program which allows the user to search for specific target words in a corpus, provide exhaustive lists for the occurrences of the word in context, enable the analysis of lexical collocations, and provide frequency information.

The concordancing program is the first and most basic task after the samples that will enable anyone seeking a particular word form to find out quickly. It is essentially indexes of every word, word's position with a set of numbers indicating the sample, the section within the sample, and the paragraph and the line which the words appear. These enable the program avoiding to search through the entire corpus letter by letter for every word and it can find all the occurrences very rapidly (Landau, 2001: 333, 335).

2.4.2.2 Concordancing Software

Tribble and Jones (1990: 93) indicated that there are many concordancing programs using for analyze the corpus, i.e. Longman Mini Concordancer, Micro OPC, Microsoft Word 4 and 5, WordCruncher, WordPerfect 5.0 and WordSmith Tools. For this study, WordSmith Tools Version 3 by Mike Scott 1999 was used to analyze sports news corpus. There are two reasons why the researcher chose this computer program. Firstly, it serves this study's needs and provides a menu-driven format with all the functions covered the objectives of this study. Secondly, it is a convenient and suitable tool for vocabulary analysis.

2.4.2.3 Text Processing

The concordancing program 'WordSmith Tools' consists of six tools, i.e. WordList, Concord, Key Word, Splitter, Dual Text Aligner, and Viewer (Mctait, 1998: Online). For this study, WordList, Concord, and Cluster which are included in Concord, were used to analyze sports news corpus because these tools have different specific target. WordList tool was used to find out word frequency lists in terms of frequency and alphabetical order, text analysis statistics, and Key Word In Context (KWIC).

2.4.2.3.1 Text Analysis Statistics

After WordSmith Tools identified the finding of running words, word form, or word type, it can produce text analysis statistics or figure for a number of relationship, i.e. tokens, types and types/token ratio as an example from psychology corpus (Kaewphanngam, 2002: 54-55) in Table 2.1 below:

Table 2.1 Statistical Details of Each Sub-field in Psychology Corpus

Statistical Details	Text Files						
	Overall	CT	CA	ET	EA	ST	SA
Tokens	236,086	52,709	26,830	55,949	33,469	24,579	42,550
Types	15,703	5,543	4,081	5,140	4,871	4,633	6,443
Type/Token Ratio	6.65	10.52	15.21	9.19	14.55	18.85	15.14

CT = Counseling Psychology Texts

CA = Counseling Psychology Articles

ET = Educational Psychology Texts

EA = Educational Psychology Articles

ST = Social Psychology Texts

SA = Social Psychology Articles

2.4.2.3.2 Word Frequency Lists

Word frequency list is the list of vocabulary in the text files. The information that is available in the frequency list can be rendered in several ways. There are two formats, i.e. frequency of occurrence (frequency order) and alphabetical order with the number of time, and line number to indicate their place in the text corpus (Kaye, 1988b: 150). Word frequency list in frequency order begins with the most frequent item and continues down to the single occurrences. The example of the frequency order of the whole corpus in the BNC (Leech et al., 2001: 120) was given in Table 2.2.

Table 2.2 Word Frequency List of the Whole Corpus in the BNC

No.	Word	Freq.	Percent
1	the	61,847	6.18
2	of	29,391	2.94
3	and	26,817	2.68
4	to	25,627	2.56

Table 2.2 (Continued)

No.	Word	Freq.	Percent
5	a	21,626	2.16
6	in	18,214	1.82
7	it	10,875	1.09
8	is	9,982	1.00
9	was	9,236	0.92
10	I	8,875	0.89

2.4.2.3.3 Key Word In Context (KWIC)

Key Word In Context is a popular type of computer-generated concordance which is easy for a researcher to scan quickly. Each line of concordance contains an instance of a select word and the page is aligned centrally around this word. The text before and after the selected word is printed, then the space, the word in question, more text, and the end of the line is reached. The full KWIC format prints a whole line of text with the word under examination in the middle (Sinclair, 1991: 32-33, 173) as an example from the sub corpus of COBUILD headword list (Renouf: 1984) shown in Figure 2.1:

r theft will be falling off markedly. Another **factor** ensuring a reduction in road-based crime further back in time, it is undeniably a major **factor** in determining design and marketing production that was brought up the most important **factor** is that life now can be prolonged and be d
sity to see what's inside an object. The main **factor** that should concern us is not the actual des
" The radical alternatives share one common **factor**. They require a general willingness on the
s and social investment of the State? A third **factor** was the way in which the growing public s
ver programmes. (Um) for example, the first **factor** we can look at would be that of purpose.
ent of the stored program is the major single **factor** which allowed computers to advance way

Figure 2.1 Concordance of 'Factor'

Having surveyed the background of the sports news, an overall theoretical framework of vocabulary and sentence structures, corpus study, and the relevance of the objectives of this

study, we begin to set the theoretical framework of the study. The next task is to present research methodology in the next chapter.



This material is reserved for educational use only, not allowed for commercial use.

Forbidden to modify the content, and cite the document when use.

CHAPTER 3

RESEARCH METHODOLOGY

The aim of this chapter is to present the research methodology used in this study. The chapter is divided into two main sections: description of the Bangkok Post and process of the study.

3.1 Description of the Bangkok Post

The sports news in the Bangkok Post was selected as the source of the sports news stories. The reasons for the selection are that the Bangkok Post is the oldest English newspaper in Thailand and has been considered the most credible and dependable for national and international news (Lent, 1982: 322). The first issue of the Bangkok Post was published on August 1, 1946 and it has been well known among English language readers in Thailand since then (Na Bangchang, 1996: 73). At present, the Bangkok Post has a circulation of 65,000 copies per day (AP, 2004: Online).

3.2 Process of the Study

The process of the study is divided into four main stages: data collection, content analysis, results and discussions, and conclusions and suggestions.

3.2.1 Data Collection

As mentioned in Chapter 1, the sports news stories of 5,375 were collected from the Bangkok Post during a six-month period of time from July 1, 2002 to December 31, 2002. They were grouped into seven categories, i.e. advance stories, following event stories, sports activity schedules, scores and standings, journalist comments, reader opinions, and interviewing articles. The sports news stories in each category were calculated into percentages. From the percentage of the collection in each category, two sports news stories were daily selected within the six-month period of time (180 days). The total number of the selected sports news stories for the content analysis was 360 (180 days x 2 news stories per day). The number of sports news stories in each category was shown in Table 1.1 in Chapter 1.

3.2.2 Content Analysis

The content analysis was divided into two parts: vocabulary analysis and sentence structure analysis.

3.2.2.1 Vocabulary Analysis

After 360 sports news stories were selected for analysis, all of them were entered into the database by scanning. At the scanning stage, Optical Character Recognition (OCR) was used to scan the 360 sports news stories. The data were stored as plain-text files (*.txt). The seven files were named as follows:

- AS.txt (Advance Stories Text File)
- FES.txt (Following Event Stories Text File)
- SaS.txt (Scores and Standings Text File)
- SAS.txt (Sports Activity Schedules Text File)
- JC.txt (Journalist Comments Text File)
- RO.txt (Reader Opinions Text File)
- IA.txt (Interviewing Articles Text File)

Each text file was exported into word document file (*.doc). The vocabulary was analyzed by using the concordancing software *'WordSmith Tool Version 3.'* The frequency of occurrence of the token (running words) and word types, and type/token ratio were calculated. In this study, vocabulary analysis was divided into three types, i.e. function words (prepositions and conjunctions), content words (nouns, verbs, adjectives, and adverbs), and abbreviations and acronyms. The results are shown in the form of frequency and percentage.

3.2.2.2 Sentence Structure Analysis

All sentences in the sports news corpus were classified manually into four main types, namely simple, compound, complex, and compound-complex sentences. The guidelines on the sentence structure analysis were based on *'A University Grammar of English'* written by Quirk and Greenbaum (1973) and *'Modern English: A Practical Reference Guide'* written by Marcella Frank (1972).

3.2.3 Results and Discussions

The findings were shown and discussed according to the objectives of this study. The results of vocabulary analysis were presented in the form of frequency and percentage. The occurrence frequencies of sentence structures were calculated and the examples of the sentences were given. The results were then followed by discussions.

3.2.4 Conclusions and Suggestions

The conclusions and suggestions were presented in Chapter 5.



CHAPTER 4

RESULTS AND DISCUSSIONS

This chapter focuses on the results and discussions of vocabulary and sentence structure analysis. The first section shows the statistical analysis of sports news corpus. The second section presents vocabulary analysis in terms of function words, content words, and abbreviations and acronyms. The third section presents the analysis of sentence structures, i.e. simple, compound, complex, and compound-complex sentences. Then, the findings will be discussed according to the objectives of the study.

4.1. Statistical Analysis of Sports News Corpus

The statistical details show the text statistics of the contents of sports news corpus. Table 4.1 summarises the frequency of occurrence of the tokens (running words), word types, type/token ratio, and hapax legomena. As mentioned in Chapter 1, the tokens are the number of all word-forms which occur in the text. These word-forms are counted each time they occur. A token which occurred more than once in different positions is counted as one word type. The type/token ratio is calculated to show how many times each word type occurs in the corpus. Hapax legomena, the words which occur only once in the corpus, are also shown.

Table 4.1 Statistical Details of Sports News Corpus

Statistical Details	Categories of Sports News							
	Overall	AS	FES	SaS	SAS	JC	RO	IA
Tokens	127,134	26,486	66,366	20,571	1,221	4,856	4,671	2,963
Word Types	12,082	4,501	7,588	3,156	245	1,483	1,387	980
Type/Token Ratio	1:11	1:6	1:9	1:7	1:5	1:3	1:3	1:3
Hapax Legomena	5,544	2,294	3,581	1,717	97	921	869	616

AS = Advance Stories

FES = Following Event Stories

SaS = Scores and Standings

SAS = Sports Activity Schedules

JC = Journalist Comments

RO = Reader Opinions

IA = Interviewing Articles

This material is reserved for educational use only, not allowed for commercial use.

Forbidden to modify the content, and cite the document when use.

From Table 4.1, the whole corpus selected from the 360 sports news stories consists of 127,134 tokens. The occurrence of the word types depends on the tokens which occurred more than once in different positions. For the whole corpus, the ratio of a word type to tokens is 1:11. This means that each word repeated 11 times on average throughout the sports news corpus. There are 5,544 word types that occur once in the sports news corpus. These words are called '*hapax legomenon*'. They are proper names, such as countries, cities, sports, teams, and players.

4.2 Vocabulary Analysis

The analysis of vocabulary was divided into three parts: function words, content words, and abbreviations and acronyms.

4.2.1 Function Words

A word frequency list is a list of all word types in a corpus. The list begins with the most frequent item (see Appendix A). Within the sports news corpus, function words, consisting of determiners, prepositions, and conjunctions, occur at the high frequency. The word '*the*' appeared at the highest rank (6,553 times or 5.15 %). This finding confirms the view of Nation (2001) that the most frequent words are function words such as '*the*', '*in*', and '*for*'. Jurafsky and Martin (2002: 314) stated that the word '*the*' is the most frequent word in most English corpora. It occurs in the noun phrases such as '*the first time*', '*the World Cup*', '*the third round*', '*the second half*'. Table 4.2 shows the top ten words in the sports news corpus, compared with the other five corpora, i.e. BNC, Birmingham Corpus, Brown Corpus, LOB Corpus, and LLC.

Table 4.2 Top 10 Words in Sports News Corpus Compared with the Other Five Corpora

No.	Sports News Corpus	BNC	Birmingham Corpus	Brown Corpus	LOB Corpus	LLC
1	the	the	the	the	the	the
2	in	of	of	of	of	I
3	to	and	and	and	and	and
4	a	to	to	to	to	to
5	and	a	a	a	a	of
6	of	in	in	in	in	a
7	's	it	that	that	that	you

Table 4.2 (Continued)

No.	Sports News Corpus	BNC	Birmingham Corpus	Brown Corpus	LOB Corpus	LLC
8	for	is	I	is	is	that
9	on	was	it	was	was	in
10	was	I	was	he	it	it

Function words come out top of the frequency list in the sports news corpus, i.e. *'the'*, *'in'*, *'to'*, *'a'*, and *'of'*. Those ranked most highly are exactly like those of the other five corpora, namely BNC, Birmingham, Brown, LOB, and LLC. The word *'the'* appeared at the highest rank in each corpus. From the top ten words in these six corpora, all of them were function words while the content words did not occur. As claimed by Smith (1991: 82), English rely on function words to serve purposes that recur in nearly every sentence. Moreover, it can be seen that *'s'* appeared at the seventh rank of the sports news corpus. It is the genitive ending *'s'*. It is counted as an enclitic particle and as a separate word token (Leech et al., 2001: 13). This shows that *'s'* associated more with informal writing like in the sports news. The sports news writers find it convenient to use the shortened form of words.

In addition, prepositions and conjunctions are function words. Table 4.3 and 4.4 show the top 14 prepositions and conjunctions occurring in the sports news corpus.

Table 4.3 Top 14 Prepositions

No	Rank	Preposition	Freq.	Percent
1	2	in	2,567	2.02
2	3	to	2,386	1.88
3	6	of	2,012	1.58
4	8	for	1,112	0.87
5	9	on	1,039	0.82
6	11	with	738	0.58
7	12	at	722	0.57
8	22	from	467	0.37
9	24	by	415	0.33

Table 4.4 Top 14 Conjunctions

No	Rank	Conjunction	Freq.	Percent
1	5	and	2,124	1.67
2	18	as	523	0.41
3	23	but	438	0.34
4	32	after	364	0.29
5	60	when	230	0.18
6	97	while	141	0.11
7	100	before	135	0.11
8	129	so	108	0.08
9	133	if	104	0.08

This material is reserved for educational use only, not allowed for commercial use.

Forbidden to modify the content, and cite the document when use.

Table 4.3 (Continued)

No	Rank	Preposition	Freq.	Percent
10	41	out	279	0.22
11	42	over	279	0.22
12	111	into	122	0.01
13	211	behind	69	0.05
14	227	through	63	0.05

Table 4.4 (Continued)

No	Rank	Conjunction	Freq.	Percent
10	150	or	94	0.07
11	171	since	84	0.07
12	186	because	76	0.06
13	470	although	32	0.03
14	750	until	20	0.02

From the word frequency list in the whole corpus, the total number of function words is 36. There are 21 prepositions (58.33 %) and 15 conjunctions (41.67 %). Table 4.3 shows that the preposition with the highest frequency is the word 'in' (2,567 times or 2.02 %). According to Kennedy (1998: 139), a preposition occurs once in every eight words in almost any English text. For conjunctions in Table 4.4, 'and' occurs at the highest rank (2,124 times or 1.67 %) in the sports news corpus. The result supports Jurafsky and Martin's (2002: 315) finding that the word 'and' is the most frequent conjunction in most English corpora and placed at the high rank.

As 'and' occurs at the highest rank in the sports news corpus, concordance lines of this word are selected to show its circumstance.

Harrington had 17 points and eight rebounds	and	Ron Mercer collected 14 points. Reggie Mil
s 102-94. Lorenzen Wright added 17 points	and	11 rebounds for Memphis, which was missi
he Wimbledon Championships. Kournikova	and	Bjorkman fought back from a set down to ta
dmitted. "They have lost a couple of players	and	I think they have also had a few injuries but
n a demonstration sport in the Asian Games	and	SEA Games but has failed to become a med
and 10 rebounds and Elton Brand added 17	and	11 as the Los Angeles Clippers posted a 99-

Figure 4.1 Concordance of 'and'

In Figure 4.1, 'and' is used to link words such as "Kournikova and Bjorkman", phrases such as "the Asian Games and Sea Games", and clauses such as "Harrington had 17 points and eight rebounds and Ron Mercer collected 14 points".

4.2.2 Content Words

The content words are open class words which are classified into nouns, verbs, adjectives and adverbs (Jurafsky and Martin, 2002: 313). The results are divided into two parts: the top twenty-eight frequent words, and the content words in each part of speech.

4.2.2.1 Top Twenty-Eight Content Words

Top twenty-eight content words in the sports news corpus, compared with the BNC are shown in Table 4.5.

Table 4.5 Top Twenty-Eight Content Words in Sports News Corpus Compared with the BNC

No.	Sports News Corpus				BNC			
	Type	Rank	Freq.	Percent	Type	Rank	Freq.	Percent
1	said	20	508	0.40	said	53	2,087	0.21
2	year	27	404	0.32	time	69	1,542	0.15
3	world	33	364	0.29	people	83	1,241	0.12
4	win	38	300	0.24	know	84	1,233	0.12
5	beat	40	294	0.23	see	87	1,186	0.12
6	game	47	268	0.21	new	89	1,145	0.11
7	new	48	267	0.21	well	91	1,119	0.11
8	team	49	264	0.21	get	97	995	0.10
9	season	50	259	0.20	way	98	958	0.10
10	points	52	258	0.20	made	101	943	0.09
11	final	53	252	0.20	got	102	932	0.09
12	time	55	249	0.20	think	104	916	0.09
13	Australia	61	228	0.18	years	107	902	0.09
14	won	63	224	0.18	go	111	881	0.09
15	round	64	208	0.16	good	118	795	0.08
16	England	65	207	0.16	make	120	791	0.08
17	Cup	66	205	0.16	year	121	737	0.07
18	match	68	203	0.16	take	127	715	0.07
19	League	70	201	0.16	own	123	695	0.07
20	games	74	184	0.14	come	124	695	0.07

Table 4.5 (Continued)

No.	Sports News Corpus				BNC			
	Type	Rank	Freq.	Percent	Type	Rank	Freq.	Percent
21	Thailand	76	182	0.14	say	125	679	0.07
22	open	80	166	0.13	going	127	658	0.07
23	victory	82	162	0.13	work	129	653	0.07
24	Reuters	84	160	0.13	government	133	633	0.06
25	top	87	152	0.12	man	135	614	0.06
26	day	88	150	0.12	day	137	610	0.06
27	old	91	149	0.12	put	140	596	0.06
28	south	92	148	0.12	world	141	590	0.06

As stated in Chapter 1, the BNC is the 100-million-word corpus containing spoken (10 million words) and written (90 million words) British English that was created by Geoffrey Leech in 1995. Leech et al. (2001) had similar findings: words such as *'said'*, *'year'*, *'world'*, *'new'*, *'time'*, and *'day'* appeared at the top 28 content words.

From Table 4.5, it can be seen that the word *'said'* appeared at the highest rank in both corpus, 508 times in sports news corpus and 2,087 times in BNC. In the sports news corpus, the word *'said'* is the past tense of *'say'* for direct speech. It is used for giving the actual words that someone uses. The following sentence is the example of the word *'said'* in the sports news stories.

- (1) It is a pleasure for us to be here in Monza for the Italian Grand Prix where our fans have given us a really warm welcome, *said* Ferrari boss Jean Todt.

Biber et al. (1999: 373, 375) stated that *'said'* is the most frequent lexical verb occurring in fictions and news. As claimed by Thompson (1991), the language used in newspapers was constructed for oral words such as *'said'*.

Unlike the BNC, the sports news stories consist of specific contents. Vocabulary is mostly related to sports. As can be seen in Table 4.5, content words appeared at the high frequency concerning sports such as *'win'*, *'beat'*, *'world'*, *'game'*, *'team'*, *'season'*. For the BNC, words with the high frequency like *'said'*, *'time'*, *'people'*, *'know'*, *'see'*, *'new'*, or *'well'* are general.

4.2.2.2 Nouns, Verbs, Adjectives, Adverbs

Content words in sports news corpus are grouped into nouns, verbs, adjectives, and adverbs, as illustrated in Table 4.6 (see Appendix C for more details).

Table 4.6 Content Words in Sports News Corpus

Noun	Freq.	Verb	Freq.	Adjective	Freq.	Adverb	Freq.
1. year	404	1. said	508	1. new	267	1. yesterday	197
2. world	364	2. beat	294	2. top	152	2. just	167
3. game	268	3. won	224	3. old	149	3. also	160
4. team	262	4. scored	133	4. south	148	4. only	147
5. season	259	5. win	123	5. Thai	129	5. now	100
6. points	258	6. played	117	6. next	118	6. well	94
7. time	249	7. play	108	7. American	111	7. here	87
8. Australia	228	8. made	107	8. national	97	8. still	69
9. round	208	9. added	94	9. good	90	9. then	65
10. England	207	10. lost	85	10. Asian	89	10. today	63
11. Cup	205	11. hit	79	11. grand	85	11. again	58
12. League	201	12. get	78	12. international	78	12. away	50
13. match	199	13. take	78	13. former	74	13. really	42
14. final	198	14. do	75	14. best	71	14. never	41
15. games	184	15. got	72	15. Australian	67	15. earlier	36
16. Thailand	182	16. going	69	16. Major	67	16. early	26
17. win	177	17. took	66	17. big	57		
18. Open	163	18. come	65	18. English	55		
19. victory	162	19. go	65	19. long	55		
20. Reuters	160	20. playing	64	20. final	54		

Table 4.6 shows the top twenty nouns, verbs, adjectives, and top 16 adverbs. From the top 459 words in the whole corpus, the total number of content words was 240. They were divided into nouns (145 words), verbs (48 words), adjectives (31 words), and adverbs (16 words) (see Appendix C). There are 39 proper names in the top 145 nouns. Most of them are the name of

countries such as 'Australia', 'England', 'Thailand', 'France', 'Germany' and so on. As mentioned in 4.1.3.1, the word 'said' is the most frequent verb found in the corpus. It is a reporting verb to describe what someone says. The verbs such as 'beat', 'won' and 'play' are words concerned with what people do in sports and what happens to them. All adjectives occurred in the sports news corpus are descriptive adjectives. They indicate inherent qualities or physical states, such as 'new', 'old', 'good', 'big', and proper adjectives, namely 'American', 'Asian', 'Australian', 'English', and 'European'. For adverbs, the word 'yesterday' (freq. 197) is the most frequent adverb in the sports news corpus. 'Yesterday' is an adverb of time showing the recent past in the sports news.

As can be seen from Table 4.5, some words, i.e. 'final' and 'win', function more than one part of speech. In the sports news corpus, the word 'final' can be identified as a noun (198 times) and an adjective (54 times). The examples are shown in the following sentences.

- (2) The **final** between the winners of Group A and B is scheduled for January 4.
- (3) Thailand kept up the pressure and Anucha made it 3-2 and then clinched his hat-trick in the **final** minute for a 4-2 scoreline.

There are interesting differences in the distribution of words across the seven categories. Table 4.7 shows the top 30 words in each category of sports news corpus (see Appendix B for more details).

Table 4.7 Top 30 Words in Each Category of Sports News Corpus

No.	AS		FES		SaS		SAS		JC		RO		IA	
	Word	Freq.	Word	Freq.	Word	Freq.	Word	Freq.	Word	Freq.	Word	Freq.	Word	Freq.
1	the	1,568	the	4,049	beat	190	V	38	the	286	the	329	the	199
2	to	640	in	1,617	the	120	ESPN	35	to	115	to	134	to	99
3	in	613	a	1,511	on	117	golf	27	a	109	and	110	in	66
4	a	537	to	1,380	Spain	107	STAR	27	of	108	of	92	of	65
5	and	518	and	1,237	results	99	Supersport	24	in	104	a	76	and	64
6	of	499	of	1,150	of	98	day	23	and	99	in	75	a	55
7	's	332	's	726	and	92	delayed	23	for	64	it	68	that	37
8	for	304	for	665	Australia	92	soccer	23	is	60	is	54	is	36
9	on	244	on	580	in	92	a.m.	21	that	57	that	54	's	32
10	will	226	was	518	division	88	Tour	20	at	50	's	53	he	29
11	said	214	with	476	m	88	PGA	19	was	47	I	51	will	27
12	is	200	his	426	at	87	baseball	18	are	45	for	50	it	25
13	he	199	at	385	League	86	Championship	18	's	43	be	40	was	25

Table 4.7 (Continued)

No.	AS		FES		SaS		SAS		JC		RO		IA	
	Word	Freq.	Word	Freq.	Word	Freq.	Word	Freq.	Word	Freq.	Word	Freq.	Word	Freq.
14	be	184	I	354	France	68	channel	18	they	37	have	39	for	23
15	with	180	first	352	England	66	League	17	with	37	on	39	his	22
16	at	176	he	324	seconds	65	Major	17	as	36	as	37	as	21
17	that	167	as	307	final	57	on	16	he	31	world	37	be	21
18	I	165	had	279	a	56	TV	16	be	30	football	32	has	20
19	has	161	it	265	Italy	55	p.m.	14	it	30	you	31	I	20
20	his	151	said	264	Germany	51	today	14	from	28	not	30	on	20
21	year	151	but	260	men	51	Open	13	have	28	n't	29	from	19
22	was	146	after	254	metres	51	tennis	12	not	28	this	29	Thailand	19
23	it	144	from	240	women	50	Thailand	12	their	28	they	28	have	18
24	have	143	win	240	bt	46	ATP	9	Thai	27	we	27	Paradorn	17
25	from	123	by	239	NY	45	European	8	but	25	game	26	world	17
26	as	119	two	239	from	44	morning	8	games	25	with	26	said	16
27	but	119	who	239	standings	43	Series	8	this	25	are	25	with	16
28	who	107	second	230	Russia	42	Masters	7	which	25	by	25	but	14
29	team	101	over	229	Spa	41	world	7	who	24	what	24	skating	14
30	this	101	that	228	overs	40	cycling	6	his	23	would	23	they	14

AS = Advance Stories

FES = Following Event Stories

SaS = Scores and Standings

SAS = Sports Activity Schedules

JC = Journalist Comments

RO = Reader Opinions

IA = Interviewing Articles

Table 4.7 shows the distinctive vocabularies of each category. In Advance Stories (AS) and Following Event Stories (FES), the word *'said'* (past tense of *'say'*) is the first content word occurring at the eleventh and twentieth rank respectively. *'Say'* is the most frequent lexical verb in news (Biber et al., 1999: 375). The following sentence is the example of using the word *'said'* in the sports news stories.

(4) Widjaja **said** her aim next season was to move into the top 50 rankings.

In Journalist Comments (JC), the first content word is *'Thai'*. For Interviewing Articles (IA), the word *'Thailand'* is the first content word occurring at the twenty-second rank, followed by *'Paradorn'*, *'world'*, *'said'*, and *'skating'*. The example of using the word *'Thai'* is shown in the following sentence.

- (5) As for **Muay Thai**, if Thai fighters win every time, it becomes boring - like Formula One at the moment which is being monopolised by Ferrari and Michael Schumacher.

In Reader Opinions (RO), the words *'world'*, *'football'*, and *'game'* occur at the seventeenth, eighteenth, and twenty-fifth ranks, respectively. This explains that readers asked various problems and expressed their opinions about World Cup 2002, as shown in (6).

- (6) Let me commend you for the superb **World Cup Special** supplement with multi-faceted presentations

Unlike the other categories, Scores and Standings (SaS) and Sports Activity Schedules (SAS), content words i.e. *'beat'* and *'V'* (Versus) were found in the first rank. The words that reflect this category are *'result'* and *'standings'* occurring at the fifth and twenty-seventh ranks, respectively. As can be seen from the top 30 words, proper names, especially countries, i.e. *'Spain'*, *'Australia'*, *'France'*, *'England'*, *'Italy'*, *'Germany'*, and *'Russia'* were frequently found in this category. This indicates that Scores and Standings (SaS) concerns a number of points and goals made by a player or team in a competition. The following is the example of using the word *'beat'* in the sports news stories.

- (7) Felix Mantilla (Spain) **beat** 7-Younes El Aynaoui (Morocco) 7-6 (7-2) 7-5.

In Sports Activity Schedules (SAS), *'V'* (Versus) was frequently presented because of the notification of schedules in various sports. The channels of TV programs of the competition in various sports were found, i.e. *'ESPN'*, *'STAR'*, and *'SuperSport'*. The names of various sports also occur in this category, i.e. *'golf'*, *'soccer'*, *'baseball'*, *'tennis'*, and *'cycling'*. In addition, the nouns of the sport tournament were also found. They are *'Tour'*, *'Championship'*, *'League'*, *'Major'*, *'Open'*, *'Series'*, and *'Masters'*. The example of using the word *'Tour'* in the sports news stories is shown in the following sentence.

- (8) 13:00 Golf: **PGA Tour Buick Open Day 4** (STAR/delayed).

4.2.3 Abbreviations and Acronyms

Abbreviations and acronyms take 1.64 % or 198 word types of the entire sports news corpus (see Appendix D for full list). The top ten abbreviations in the corpus are as follows:

This material is reserved for educational use only, not allowed for commercial use.

Forbidden to modify the content, and cite the document when use.

Table 4.8 Top 10 Abbreviations in Sports News Corpus

No.	Abbreviations	Word	Examples	Freq.
1	's	is, has	It's not easy to face a player such as Safin on hard-court.	1,415
2	n't	not	I haven't played in a tournament for months and this will be good preparation for the next English season.	327
3	No.	Number	Yevgeny Kafelnikov and world No. 3 Marat Safin appeared well set for victory after taking the third set and leading 2-0 early in the fourth.	274
4	US	United States	Woods, 26, was a five-times winner on the PGA Tour this year, including the US Masters and US Open.	226
5	St.	Saint, State	In Nashville, Brent Johnson made 35 saves in his first appearance of the season as the St. Louis Blues battled to a 2-2 tie with the Predators.	115
6	AFP	Associated French Press	Toronto, AFP Cristiano da Matta is on a roll in the Championship Auto Racing Teams (CART) series.	80
7	m	metre	He also swam on the USA's world record-setting 4x100 m medley relay.	67

Table 4.8 (Continued)

No.	Abbreviations	Word	Examples	Freq.
8	a.m.	ante meridiem	The quarter-final round matches are scheduled to be played today at the same venue, with the matches starting at 10 a.m.	66
9	'm	am	I didn't win the French Open, but I'm going to do my best at Wimbledon.	57
10	've	have	I've won my last four finals and I'd like to continue that record.	57

According to Ajjan (1993: x), abbreviation is any shortened form of a word or phrase used to present the complete form. From Table 4.8, the abbreviation 's' has the highest frequency (1,415 times), followed by 'n't' (327 times), 'No.' (274 times), 'US' (226 times), and 'St.' (115 times), respectively. There are four contractions found in the top ten abbreviations, i.e. 's', 'n't', 'm', and 've'.

In this study, both abbreviations and acronyms are found (see Appendix D for full list). As can be seen from Table 4.8, all words at the top ten ranks were abbreviations, whereas acronyms did not occur. Table 4.9 shows the examples of abbreviations and acronyms found in the sports news stories.

Table 4.9 Examples of Abbreviations and Acronyms in Sports News Corpus

Type	Freq.	Percent	Abbreviations	Word
Abbreviations	4,092	98.67	's	is, has (freq. 1,415)
			n't	not (freq. 327)
			No.	Number (freq. 274)
			US	United States (freq. 226)
			St.	Saint, State (freq. 115)

Table 4.9 (Continued)

Type	Freq.	Percent	Acronyms	Word
Acronyms	55	1.33	FIFA	Federations of International Football Association (freq. 12)
			CART	Championship Auto Racing Teams (freq. 8)
			TALGA	Thailand Amateur Ladies Golf Association (freq. 7)
			ASEAN	Association of South-East Asian Nations (freq. 4)
			FIBA	Federation of International Basketball Association (freq. 3)
Total	4,147	100		

If the overall abbreviations are closely considered, it is found that there were 4,092 abbreviations (98.67 %) and 55 acronyms (1.33 %). This indicates that abbreviations are used more than acronyms in the sports news. As can be seen from the first five ranks of acronyms, 4 of them are athletic organizations, i.e. 'FIFA', 'CART', 'TALGA', and 'FIBA'. The use of abbreviations and acronyms in the sports news corpus indicates that the sports news writers find it convenient to shorten the words into abbreviations because they can save time for writing and save space for printing.

To summarise, the purpose of this study is to classify vocabulary into function words, content words, and abbreviations and acronyms. The number of vocabulary in each type is presented in Table 4.10.

Table 4.10 Number of Vocabulary in Sports News Corpus

Vocabulary	Word Types	Percentage
Function Words	60	0.50
Content Words	11,824	97.86
Abbreviations and Acronyms	198	1.64
Total	12,082	100.00

According to the Table 4.10, the total word types of sports news stories were 12,082. Content words occurred with the highest number (11,824 word types or 97.86 %), followed by abbreviations and acronyms (198 word types or 1.64 %), and function words (60 word types or 0.50 %), respectively.

4.3 Sentence Structure Analysis

The selected data consists of 3,612 sentences which include four main sentence structures: simple, compound, complex and compound-complex sentences. The frequency of the occurrence of the sentence structures are shown in Table 4.11.

Table 4.11 Frequency of Occurrence of Sentence Structures

Sentence Structures	Frequency of Occurrence	Percentage
Simple Sentences	1,215	33.64
Compound Sentences	122	3.38
Complex Sentences	2,029	56.17
Compound-Complex Sentences	246	6.81
Total	3,612	100.00

As shown in Table 4.11, of the 3,612 sentences, complex sentences occurred with the highest frequency (56.17 %), followed by simple sentences (33.64 %), compound-complex sentences (6.81 %), and compound sentences (3.38 %). The result indicates that in this study, sports journalists mostly used the complex sentences. As claimed by Booher and Hill (1989: 50), the complex sentences add variety to writing and help emphasize the most important ideas. The complex sentences are used when journalists would like to emphasize one idea over another in a sentence (Langan, 1993: 392). The result, moreover, supports Booher and Hill's statement that to get to the point clearly, quickly, and emphatically, journalists or reporters prefer to use the complex and simple sentences in reporting sports news. Simple sentences are also recommended because they are direct, and easy to read and understand (Booher and Hill, 1989: 49). In addition, long sentences (20 words or up) slow down communication and delay the completion of the thought with other clauses or phrases (Roze, 1997: 15). Accordingly, journalists may avoid using compound-complex sentences. As compound sentences which ranked fourth have two or more

independent clauses, they divide readers' attention and force consideration of two or more ideas at once (Booher and hill, 1989: 50). The readers, therefore, must slow their pace to concentrate on both clauses. As a result, journalists may not prefer compound sentences.

The sentence structures found in the sports news stories are concluded below. Some examples are shown in Appendix E.

4.3.1 Simple Sentences

The following are the examples of simple sentences found in the sports news stories.

- (1) He played a great match.
- (2) Only five women's teams will compete.
- (3) This course is a fair test.
- (4) Only one main draw match was on the opening day programme.
- (5) They lost their Tri-Nations title to New Zealand.
- (6) The scorers gave Harrison a 59-55 victory.
- (7) The Broncos made it respectable with three second-half touchdowns.
- (8) There aren't many players like Figo, Zidane, Raul, Roberto Carlos, Fernando Hierro, Fernando Morientes, Ivan Helguera.

For SVA and SVOA, as mentioned in Chapter 2, an adverbial (A) is obligatory in the sense that it is required for the complementation of the verb. That is, the sentence is incomplete if the element is omitted, as shown in (9) and (10).

- (9) Bettis and Sapp were among the pre-game combatants.

S V A

- (10) Martin Lapointe's second period goal tied the game for Boston.

S V O A

An adverbial may occur in the form of an optional adverbial which may be added to or removed from a sentence without affecting its acceptability. Moreover, it is generally mobile. That is, it is capable of occurring in more than one position in the clause, as shown in (11), (12) and (13).

(11) Only two races were completed (on Monday).

S V (A)

(12) Brazil produced their best performances (from the second round).

S V O (A)

(13) The Bombers will (now) meet Port Adelaide (next week)

S (A) V O (A)

(after the 17.9 (111) to 11.12 (78) win).

(A)

4.3.2 Compound Sentences

Compound sentences used in the sports news are illustrated in (14), (15), (16), (17), (18) and (19). As stated in Chapter 2, the compound sentences consist of two or more than two simple sentences with full predications. Two simple sentences with full predications are joined by coordinate conjunctions: *and*, *so*, or *but*, or by conjunctive adverbs: *hence*.

(14) I didn't have my form and I lost my game.

(15) Thailand have often beaten Australia in the past few years, so Shanrit is not much worried about them.

(16) The six and two-mile starts are at the same place on the track and this probably caused the jockey's confusion.

(17) Leeds United arrive in Bangkok tomorrow morning and there will be a press conference at the Merchant Court Hotel at 2 pm on the same day.

(18) Thousands of English fans have travelled to Bratislava for the game but no other incidents were reported.

(19) They have been winning trophies in previous Skate Asia competitions and this year they are well prepared, hence they should do well.

4.3.3 Complex Sentences

Complex sentences found in the sports news stories are illustrated in (20), (21), (22), (23) and (24). As mentioned in Chapter 2, in complex sentences, dependent clauses are in the form of finite or non-finite clauses. Finite clauses consist of adverbial clauses, adjective clauses, or noun clauses. Non-finite clauses consist of -ing participle, -ed participle, to-infinitive, verbless clause, or bare infinitive.

This material is reserved for educational use only, not allowed for commercial use.

Forbidden to modify the content, and cite the document when use.

(20) I don't think *Ronaldo has been the best player in Europe.*

Independent Clause *Dependent Clause (Noun Clause)*

Clause

(21) Thongchai Jaidee hit a four-under 68 *while Prayad Marksaeng was two shots*

Independent Clause *Dependent Clause (Adverbial Clause)*

further behind with a 70.

(22) Dirk Nowitzki scored 30 points *to lead Germany to an 88-76 win over*

Independent Clause *Dependent Clause (to-infinitive)*

China.

(23) Third-seeded Russian Yevgeny Kafelnikov was also successful in his second

Independent Clause

round match, *defeating Denmark's Kenneth Carlsen 3-6, 6-4, 6-4.*

Dependent Clause (-ing participle)

(24) There's nothing wrong with the attention *she's brought to the game.*

Independent Clause

Dependent Clause

(Adjective Clause)

Moreover, some structures were found in the sports news stories. The examples of the postponement and cleft sentences are shown in (25) and (26).

(25) Postponement

It is not possible for me *to win this stage.*

(cf: To win this stage is not possible for me.)

(26) Cleft Sentences

It was *Bari* who increased their lead three minutes into injury time.

(cf: Bari increased their lead three minutes into injury time.)

4.3.4 Compound-Complex Sentences

A compound-complex structure has two independent clauses and one or more dependent clauses. Dependent clauses are in the form of finite or non-finite clauses. The following are the examples of compound-complex sentences.

CHAPTER 5

CONCLUSIONS AND SUGGESTIONS

This chapter presents the conclusions and suggestions of the study. The first part concerns the conclusions of this thesis. The second part focuses on the implications of the study.

5.1 Conclusions

The purpose of the study is to analyze vocabulary and sentence structures used in the sports news in the Bangkok Post. Five thousand three hundred and seventy-five sports news stories were collected from the Bangkok Post during a six-month period of time. The 5,375 sports news stories were grouped into seven categories, namely advance stories, following event stories, sports activity schedules, scores and standings, journalist comments, reader opinions, and interviewing articles. Three hundred and sixty sports news stories used for the content analysis were randomly selected from the 5,375 news stories. Research tools used in this study was the concordancing software 'WordSmith Tools Version 3.' The 360 printed sports news stories were scanned and stored as plain-text files (*.txt) in the corpus by using Optical Character Recognition (OCR). Each text file was exported into word document file (*.doc).

The vocabulary was analyzed by using the concordancing software 'WordSmith Tool Version 3'. The Wordlist tool was used to calculate the frequency of occurrence of the tokens, word types, and type/token ratio. The findings of word frequency lists were categorized into function words (prepositions and conjunctions), content words (nouns, verbs, adjectives, and adverbs), and abbreviations and acronyms. The results of the vocabulary analysis were presented in the form of word frequency and percentage. In this thesis, the sentence structure analysis was done manually. The guidelines on the analysis were based on '*A University Grammar of English*' written by Quirk and Greenbaum (1973) and '*Modern English: A Practical Reference Guide*' written by Marcella Frank (1972). The sentences in the sports news stories were identified and grouped into four main types, namely simple, compound, complex, and compound-complex sentences. The frequency and percentage of the occurrence of sentence structures were calculated and the examples of the sentences were shown.

According to the statistical analysis of the sports news corpus, the whole corpus selected from the 360 sports news stories consists of 127,134 tokens. The total word types of all categories

This material is reserved for educational use only, not allowed for commercial use.

Forbidden to modify the content, and cite the document when use.

of the sports news was 12,082. The ratio of word type to tokens of the whole corpus was 1:11. This means each word repeated 11 times on average throughout the sports news corpus.

In vocabulary analysis, content words occurred with the highest number of 12,082 word types (11,824 word types or 97.86 %), followed by abbreviations and acronyms (198 word types or 1.64 %), and function words (60 word types or 0.50 %), respectively. Function words such as *'the'*, *'in'*, and *'of'* occurred at the high frequency. Regarding content words, the words *'win'*, *'beat'*, *'game'*, *'points'*, *'won'*, *'match'*, *'league'*, *'victory'*, *'tournament'*, *'players'*, *'scored'*, and *'champion'* were frequently found and related to the sports matters. Proper names, i.e. the names of sports players and countries, were frequently occurred in the sports news stories. The word *'said'* appeared at the highest rank. In the sports news corpus, the word *'said'* was used to report what people said by using direct speech. In terms of prepositions and conjunctions, *'in'* and *'and'* appeared with the highest frequency. There were 198 abbreviations and acronyms in the whole corpus, which consisted of 98.67 % abbreviations and 1.33 % acronyms.

For the sentence structure analysis, of the 3,612 sentences, complex sentences occurred with the highest frequency (56.17 %), followed by simple sentences (33.64 %), compound-complex sentences (6.81 %), and compound sentences (3.38 %). Sports journalists mostly used the complex sentences. Complex sentences added variety to writing and helped emphasize the most important ideas. It was used when journalists wanted to emphasize one idea over another in a sentence. To get to the point clearly, quickly, and emphatically, journalists or reporters preferred to use the complex and simple sentences in reporting sports news. Simple sentences were also recommended because they were direct, and easy to read and understand.

5.2 Implications for Sports News Education

5.2.1 Implications for Sports Journalists

Sports journalists or sports reporters who need to write their articles in English are able to use this study as a guideline for writing the sports news stories in English daily newspapers. Such findings include high frequency words, abbreviations, and sentence structures most favorably used in the sports news.

5.2.2 Implications for Material Writers and Course Designers

The present research reveals that with concordancing software, it is easier for material writers to develop materials and provide more useful resources than in the past. The sports news

stories can be applied for material writers to help develop and evaluate vocabulary for the teaching and learning of a language. This will hopefully lead to the development of appropriate learning materials to assist students or readers in gaining greater proficiency in reading English daily newspapers.

The following sample gapfill exercise was chosen from the authentic sports news corpus. The teachers and syllabus designers may blank the selected words and ask the students to fill in the missing words as shown in the following exercise.

Use each noun once to complete these sentences.

year points one world time

1. Faldo accepts world number Tiger Woods will still be the man to beat if he wants to win a fourth Open title.
2. Hicham El Guerrouj and Paula Radcliffe were named Athletes of the by the International Association of Athletics Federations (IAAF) on Sunday.
3. Some matches will be played in the provinces for the first - in Chonburi from November 29 to December 1 and in Ayutthaya from January 11-12.
4. Jermaine O'Neal had 25 points for Indiana (14-3) while Al Harrington had 17 and eight rebounds and Ron Mercer collected 14 points.
5. The co-hosts South Korea have proved that they can play with any team in the

5.3 Suggestions for Further Studies

From the findings of the present research, further studies should be carried out to gain a deeper understanding and applications. There are three suggestions derived from this study.

5.3.1 In order to obtain an insight into language of sports news, further studies on other lexical features such as compound nouns and collocations should also be included.

5.3.2 Further research on vocabulary and sentence structures of sports news should carry out by comparing this study with sports news in other countries such as England, Australia, USA.

5.3.3 If possible, other kinds of news sections such as business, shipping, outlook, or editorial, should be studied in order to compare their linguistic features with the sports news.

BIBLIOGRAPHY

- Aiex, N. K. 1988. "Using Newspapers as Effective Teaching Tools." ERIC Digest Number 10. Indiana, U.S. ERIC No. ED300847.
- Ajjan, Diana. 1993. **21st Century Dictionary of Acronyms and Abbreviations**. New York: The Philip Lief Group, Inc.
- AP. **Bangkok Post**. [Online]. Available: <http://www.taipeitimes.com/News/World/archives>. 2004
- Arunmanakul, Virot. 2002. **Corpus Linguistics**. Bangkok: Chulalongkorn University Press.
- Biber et al. 1998. **Corpus Linguistics: Investigating Language Structure and Use**. Cambridge: Cambridge University Press.
- Biber et al. 1999. **Longman Grammar of Spoken and Written English**. London: Longman.
- Booher, D. and Hill, T. H. 1989. **Writing for Technical Professionals**. New York: John Wiley & Son, Inc.
- Boonlikhit, Phumisit. **The Analysis of Sports News Coverage in Daily Newspaper**. M.A. Thesis, Chulalongkorn University. 1998.
- Boswood, T. 1997. **New Ways of Using Computers in Language Teaching**. Illinois: Tesol inc.
- Carter, R. and McCarthy, M. 1995. "Grammar and the Spoken Language." **Applied Linguistics** 16: 141-158.
- Crutchley, R. August 18, 2002. Tosatjawong, M. Interviewer. **Classification of Sports News in English Daily Newspaper**. The Bangkok Post Building.
- Crystal, D. 1992. **An Encyclopedia Dictionary of Language and Language**. USA: Cambridge University Press.
- Frank, Marcella. 1972. **Modern English: A Practical Reference Guide**. New York: Prentice-Hall, Inc.
- Gallerstam, Martin. 1991. "Modern Swedish Text Corpora." in: **Directions in Corpus Linguistics - Proceedings of Nobel Symposium 82 Stockholm, 4-8 August**. Berlin: Mouton de Gruyter.
- Hartmann, R. R. K. 2001. **Teaching and Researching Lexicography**. England: Pearson Education Limited.
- Hunston, Susan. 2002. **Corpora in Applied Linguistics**. Cambridge: Cambridge University Press.

This material is reserved for educational use only, not allowed for commercial use.

Forbidden to modify the content, and cite the document when use.

- Iaumsupanimit, Somchit. 1986. **English in Newspapers**. Nakhon Pathom: English Department, Faculty of Arts, Silpakorn University.
- Jeffery, G. 1953. **Foreward in General Service List of English Words**. London: Longman, Gree & Co.
- Jurafsky, D. and Martin, J. H. 2002. **Speech and Language Processing: An Introduction to Natural Language Processing, Computational Linguistics, and Speech Recognition**. New Jersey: Prentice-Hall, Inc.
- Kaewphanngam, Chaiwat. **A Corpus Analysis of Psychology Texts as a Basis for the Development of Teaching Materials in English for Academic Purposes**. M.A. Thesis, Mahidol University. 2002.
- Kaye, Geoffrey. 1988b. "Computerized Corpora for the PC." Paper presented at The 9th International Conference on English Language Research on Computerized Corpora at Birmingham University, England, 18-22 May 1988.
- Kennedy, Graeme. 1998. **An Introduction to Linguistics**. London: Longman.
- Kraisornkovit, Prapai. 1997. **Redesigning the Bangkok Post to Improve its Competitive Edge in Information Society**. M.A. Thesis, Chulalongkorn University.
- Landau, S. 2001. **Dictionaries: The Art and Craft of Lexicography**. Cambridge: Cambridge University Press.
- Langan, John. 1993. **Sentence Skills: A Workbook for Writers**. New York: McGraw-Hill, Inc.
- Leech, G. 1997. **Teaching and Language Corpora: A Convergence**. In Candlin (ed.) *Teaching and Language Corpora*. London: Addison Wesley Longman Limited: 1-24.
- Leech et al. 2001. **Word Frequency in Written and Spoken English: Based on the British National Corpus**. London: Longman University Press.
- Lent, John A. 1982. **Newspaper in Asia: Contemporary Trends and Problem**. Hong Kong: Heinemann.
- Mallikamas, P. 1999. **Applications of Corpora in Language Teaching**. *Thai TESOL Bulletin*, 12 (1): 1-17.
- Mctait, K. **A Survey of Corpus Analysis Tool**. [Online]. Available: <http://lismore.ccl.umist.ac.uk/kevim/Survey.pdf>. 1998.
- Na Bangchang, Prapoj. **Factors Related to the Growth of Sports Daily Newspapers**. M.A. Thesis, Chulalongkorn University. 1996.

- Nation, I. S. P. 2001. **Learning Vocabulary in Another Language**. Cambridge: Cambridge University Press.
- Phanphanich, Prapassorn. **An Analysis of Political News in English Daily Newspaper: A Case of Bangkok Post**. M.A. Thesis, Ramkhamhaeng University. 1999.
- Poonpon, Kornwipa. **Vocabulary Input in English for Science Courses: A Corpus Analysis of Intensive and Extensive Course Materials**. M.A. Dissertation. Mahidol University. 2002.
- Quirk, R. and Greenbaum, S. 1973. **A University Grammar of English**. Essex: Longman.
- Renouf, A. J. 1984. "Corpus Development at Birmingham University" in **Corpus Linguistics: Recent Developments in the Use of Computer Corpora in English Language Research**, J. Aarts and W. Meijs (eds.). Rodopi, Amsterdam.
- Roze, Maris. 1997. **Technical Communication: The Practical Craft**. Third Edition. New Jersey: Prentice-Hall, Inc.
- Shei, C. **Collocation, Learner Corpus, Language Teaching**. [Online]. Available: <http://www.dai.ed.ac.uk/homes/shei/collocation/top.html>. 1999
- Sinclair, John. 1991. **Corpus, Concordance, Collocation**. Oxford: Oxford University Press.
- Sittirak, Nantana. **An Analytical Study of Soccer News Translation in the Star's Soccer Daily Newspaper**. M.A. Thesis, Mahidol University. 2002.
- Smith, George W. 1991. **Computer and Human Language**. New York: Oxford University Press.
- Tang-U-Thaisuk, Kanlayanee. **The Language Analysis of Information Technology News: A Case Study of the Bangkok Post and The Nation**. Thematic Paper, Ramkhamhaeng University. 2000.
- Thanomsak, Pussadee. **An Analysis of Front Page News in an English Daily Newspaper: A Case of Bangkok Post**. M.A. Thesis, Ramkhamhaeng University. 1998.
- Thompson, David R. 1991. "Oral History Lives: A Content Analysis of Newspaper Use of Language." Texas, U.S. ERIC No. ED337840.
- Tribble, C. and Jones G. 1990. **Concordance in the Classroom: A Resource book for Teachers**. Houston, TX: Athelstan.
- Walters, Timothy N. 1994. "After the Highwayman: Syntax and Successful Placement of Press Releases in Newspapers." ERIC No. EJ496063. Vol: 20; Issue: 4; Page: 345-56.



This material is reserved for educational use only, not allowed for commercial use.

Forbidden to modify the content, and cite the document when use.

APPENDIX A

Word Frequency List in Sports News Corpus

Rank	Word	Freq.	Percent
1	the	6,553	5.15
2	in	2,567	2.02
3	to	2,386	1.88
4	a	2,346	1.85
5	and	2,124	1.67
6	of	2,012	1.58
7	's	1,415	1.11
8	for	1,112	0.87
9	on	1,039	0.82
10	was	753	0.59
11	with	738	0.58
12	at	722	0.57
13	his	637	0.50
14	I	626	0.49
15	he	600	0.47
16	that	545	0.43
17	it	534	0.42
18	as	523	0.41
19	is	521	0.41
20	said	508	0.40
21	first	476	0.37
22	from	467	0.37
23	but	438	0.34
24	by	415	0.33
25	be	413	0.32
26	have	411	0.32
27	year	404	0.32
28	will	394	0.31
29	who	390	0.31
30	had	378	0.30
31	two	365	0.29
32	after	364	0.29
33	world	364	0.29
34	one	360	0.28
35	has	346	0.27
36	n't	327	0.26
37	their	307	0.24
38	win	300	0.24
39	this	297	0.23
40	beat	294	0.23
41	out	279	0.22
42	over	279	0.22
43	second	276	0.22
44	no	274	0.22
45	an	269	0.21
46	last	269	0.21
47	game	268	0.21
48	new	267	0.21
49	team	264	0.21
50	season	259	0.20
51	they	259	0.20
52	points	258	0.20

This material is reserved for educational use only, not allowed for commercial use.

Forbidden to modify the content, and cite the document when use.

Rank	Word	Freq.	Percent
53	final	252	0.20
54	are	249	0.20
55	time	249	0.20
56	we	245	0.19
57	all	243	0.19
58	three	242	0.19
59	up	237	0.19
60	when	230	0.18
61	Australia	228	0.18
62	US	226	0.18
63	won	224	0.18
64	round	208	0.16
65	England	207	0.16
66	Cup	205	0.16
67	four	203	0.16
68	match	203	0.16
69	League	201	0.16
70	yesterday	197	0.15
71	which	190	0.15
72	games	184	0.14
73	been	183	0.14
74	Thailand	182	0.14
75	against	181	0.14
76	were	180	0.14
77	just	167	0.13
78	open	166	0.13
79	third	162	0.13
80	victory	162	0.13
81	also	160	0.13
82	Reuters	160	0.13
83	would	156	0.12
84	her	154	0.12
85	top	152	0.12
86	day	150	0.12
87	off	150	0.12
88	five	149	0.12
89	old	149	0.12
90	South	148	0.12
91	only	147	0.12
92	play	146	0.11
93	back	144	0.11
94	Tournament	142	0.11
95	him	141	0.11
96	while	141	0.11
97	players	140	0.11
98	there	138	0.11
99	before	135	0.11
100	scored	133	0.10
101	division	131	0.10
102	set	131	0.10
103	Thai	129	0.10
104	home	128	0.10
105	you	128	0.10
106	Spain	126	0.10
107	champion	124	0.10
108	France	124	0.10
109	more	123	0.10
110	into	122	0.10
111	title	121	0.10

Rank	Word	Freq.	Percent
112	results	120	0.09
113	she	119	0.09
114	lead	118	0.09
115	next	118	0.09
116	run	118	0.09
117	Tour	118	0.09
118	played	117	0.09
119	women	117	0.09
120	St.	115	0.09
121	seconds	113	0.09
122	American	111	0.09
123	other	111	0.09
124	player	110	0.09
125	record	110	0.09
126	can	109	0.09
127	championship	109	0.09
128	so	108	0.08
129	under	108	0.08
130	made	107	0.08
131	my	107	0.08
132	if	104	0.08
133	than	104	0.08
134	about	103	0.08
135	down	102	0.08
136	united	102	0.08
137	fourth	101	0.08
138	minutes	101	0.08
139	now	100	0.08
140	six	100	0.08
141	winning	99	0.08
142	national	97	0.08
143	Saturday	97	0.08
144	both	96	0.08
145	finals	96	0.08
146	men	95	0.07
147	added	94	0.07
148	Germany	94	0.07
149	or	94	0.07
150	well	94	0.07
151	football	92	0.07
152	years	92	0.07
153	best	91	0.07
154	goal	91	0.07
155	place	91	0.07
156	good	90	0.07
157	Asian	89	0.07
158	could	89	0.07
159	hit	89	0.07
160	week	89	0.07
161	golf	88	0.07
162	here	87	0.07
163	quarter	87	0.07
164	Sunday	87	0.07
165	million	86	0.07
166	them	86	0.07
167	grand	85	0.07
168	like	85	0.07
169	lost	85	0.07
170	since	84	0.07

Rank	Word	Freq.	Percent
171	matches	83	0.07
172	Paradorn	83	0.07
173	eight	81	0.06
174	Italy	81	0.06
175	AFP	80	0.06
176	club	78	0.06
177	get	78	0.06
178	international	78	0.06
179	Michael	78	0.06
180	take	78	0.06
181	teams	78	0.06
182	race	77	0.06
183	tennis	77	0.06
184	today	77	0.06
185	because	76	0.06
186	metres	76	0.06
187	do	75	0.06
188	star	75	0.06
189	then	75	0.06
190	former	74	0.06
191	me	74	0.06
192	series	74	0.06
193	Williams	74	0.06
194	half	73	0.06
195	number	73	0.06
196	side	73	0.06
197	some	73	0.06
198	what	73	0.06
199	Africa	72	0.06
200	champions	72	0.06
201	got	72	0.06
202	playing	72	0.06
203	coach	71	0.06
204	going	71	0.06
205	seed	70	0.06
206	seven	70	0.06
207	behind	69	0.05
208	being	69	0.05
209	still	69	0.05
210	event	68	0.05
211	Japan	68	0.05
212	left	68	0.05
213	Australian	67	0.05
214	m	67	0.05
215	major	67	0.05
216	a.m.	66	0.05
217	most	66	0.05
218	runs	66	0.05
219	took	66	0.05
220	come	65	0.05
221	go	65	0.05
222	overs	65	0.05
223	shot	64	0.05
224	career	63	0.05
225	goals	63	0.05
226	sports	63	0.05
227	through	63	0.05
228	way	63	0.05
229	another	62	0.05

This material is reserved for educational use only, not allowed for commercial use.

Forbidden to modify the content, and cite the document when use.

Rank	Word	Freq.	Percent
230	innings	62	0.05
231	London	62	0.05
232	did	61	0.05
233	end	61	0.05
234	Friday	61	0.05
235	much	61	0.05
236	early	60	0.05
237	including	60	0.05
238	make	60	0.05
239	ball	59	0.05
240	city	59	0.05
241	Mark	59	0.05
242	think	59	0.05
243	Tuesday	59	0.05
244	very	59	0.05
245	again	58	0.05
246	David	58	0.05
247	Russia	58	0.05
248	winner	58	0.05
249	big	57	0.04
250	injury	57	0.04
251	its	57	0.04
252	'm	57	0.04
253	Monday	57	0.04
254	nine	57	0.04
255	opening	57	0.04
256	standings	57	0.04
257	start	57	0.04
258	Toronto	57	0.04
259	've	57	0.04
260	between	56	0.04
261	par	56	0.04
262	Woods	56	0.04
263	Chicago	55	0.04
264	English	55	0.04
265	fans	55	0.04
266	field	55	0.04
267	gold	55	0.04
268	long	55	0.04
269	month	55	0.04
270	yards	55	0.04
271	championships	54	0.04
272	chance	54	0.04
273	China	54	0.04
274	Florida	54	0.04
275	people	54	0.04
276	put	54	0.04
277	Brazil	53	0.04
278	European	53	0.04
279	bt	51	0.04
280	Canada	51	0.04
281	during	51	0.04
282	group	51	0.04
283	same	51	0.04
284	Texas	51	0.04
285	away	50	0.04
286	baseball	50	0.04
287	Boston	50	0.04
288	lot	50	0.04

Rank	Word	Freq.	Percent
289	point	50	0.04
290	rebounds	50	0.04
291	soccer	50	0.04
292	came	49	0.04
293	finished	49	0.04
294	leading	49	0.04
295	Masters	49	0.04
296	Philadelphia	49	0.04
297	Spa	49	0.04
298	went	49	0.04
299	Cincinnati	48	0.04
300	defeat	48	0.04
301	despite	48	0.04
302	French	48	0.04
303	Korea	48	0.04
304	man	48	0.04
305	sport	48	0.04
306	test	48	0.04
307	told	48	0.04
308	any	47	0.04
309	Atlanta	47	0.04
310	course	47	0.04
311	little	47	0.04
312	manager	47	0.04
313	many	47	0.04
314	ranked	47	0.04
315	too	47	0.04
316	Wednesday	47	0.04
317	Arizona	46	0.04
318	ATP	46	0.04
319	better	46	0.04
320	Britain	46	0.04
321	fifth	46	0.04
322	minute	46	0.04
323	Montreal	46	0.04
324	NY	46	0.04
325	past	46	0.04
326	states	46	0.04
327	times	46	0.04
328	V	46	0.04
329	Francisco	45	0.04
330	inning	45	0.04
331	PGA	45	0.04
332	should	45	0.04
333	State	45	0.04
334	Sydney	45	0.04
335	Association	44	0.03
336	Carlos	44	0.03
337	Houston	44	0.03
338	Pittsburgh	44	0.03
339	real	44	0.03
340	Tampa	44	0.03
341	total	44	0.03
342	Czech	43	0.03
343	loss	43	0.03
344	rugby	43	0.03
345	Thursday	43	0.03
346	west	43	0.03
347	Wimbledon	43	0.03

Rank	Word	Freq.	Percent
348	cricket	42	0.03
349	Detroit	42	0.03
350	Leeds	42	0.03
351	ninth	42	0.03
352	our	42	0.03
353	really	42	0.03
354	score	42	0.03
355	see	42	0.03
356	country	41	0.03
357	court	41	0.03
358	crowd	41	0.03
359	earlier	41	0.03
360	great	41	0.03
361	losing	41	0.03
362	never	41	0.03
363	night	41	0.03
364	right	41	0.03
365	scoring	41	0.03
366	where	41	0.03
367	Agassi	40	0.03
368	ago	40	0.03
369	Colorado	40	0.03
370	even	40	0.03
371	give	40	0.03
372	Johnson	40	0.03
373	leg	40	0.03
374	Republic	40	0.03
375	sixth	40	0.03
376	Chris	39	0.03
377	Cleveland	39	0.03
378	hard	39	0.03
379	India	39	0.03
380	Melbourne	39	0.03
381	result	39	0.03
382	scores	39	0.03
383	straight	39	0.03
384	bay	38	0.03
385	competition	38	0.03
386	defending	38	0.03
387	each	38	0.03
388	following	38	0.03
389	later	38	0.03
390	Minnesota	38	0.03
391	performance	38	0.03
392	USA	38	0.03
393	Argentina	37	0.03
394	British	37	0.03
395	gave	37	0.03
396	how	37	0.03
397	making	37	0.03
398	officials	37	0.03
399	once	37	0.03
400	Seattle	37	0.03
401	singles	37	0.03
402	Switzerland	37	0.03
403	tie	37	0.03
404	Western	37	0.03
405	Bangkok	36	0.03
406	Belgium	36	0.03

Rank	Word	Freq.	Percent
407	ESPN	36	0.03
408	high	36	0.03
409	hits	36	0.03
410	white	36	0.03
411	wins	36	0.03
412	Yankees	36	0.03
413	Anaheim	35	0.03
414	AP	35	0.03
415	close	35	0.03
416	face	35	0.03
417	getting	35	0.03
418	hours	35	0.03
419	Oakland	35	0.03
420	own	35	0.03
421	're	35	0.03
422	challenge	34	0.03
423	finish	34	0.03
424	having	34	0.03
425	Hewitt	34	0.03
426	know	34	0.03
427	Manchester	34	0.03
428	must	34	0.03
429	Pakistan	34	0.03
430	racers	34	0.03
431	red	34	0.03
432	stadium	34	0.03
433	stage	34	0.03
434	without	34	0.03
435	action	33	0.03
436	always	33	0.03
437	Asia	33	0.03
438	Barcelona	33	0.03
439	break	33	0.03
440	Cubs	33	0.03
441	days	33	0.03
442	doubles	33	0.03
443	held	33	0.03
444	Madrid	33	0.03
445	Martin	33	0.03
446	part	33	0.03
447	seeded	33	0.03
448	Sox	33	0.03
449	weekend	33	0.03
450	wickets	33	0.03
451	although	32	0.03
452	fight	32	0.03
453	Jones	32	0.03
454	Kansas	32	0.03
455	Kim	32	0.03
456	led	32	0.03
457	Mets	32	0.03
458	Milwaukee	32	0.03
459	president	32	0.03
460	work	31	0.02
461	ahead	31	0.02
462	athletes	31	0.02
463	captain	31	0.02
464	double	31	0.02
465	ever	31	0.02

Rank	Word	Freq.	Percent
466	help	31	0.02
467	however	31	0.02
468	looking	31	0.02
469	Milan	31	0.02
470	money	31	0.02
471	North	31	0.02
472	reach	31	0.02
473	Richard	31	0.02
474	Ronaldo	31	0.02
475	Steve	31	0.02
476	strong	31	0.02
477	Sweden	31	0.02
478	tomorrow	31	0.02
479	weeks	31	0.02
480	amateur	30	0.02
481	coming	30	0.02
482	form	30	0.02
483	John	30	0.02
484	Juan	30	0.02
485	key	30	0.02
486	meet	30	0.02
487	Miami	30	0.02
488	Paul	30	0.02
489	p.m.	30	0.02
490	Premiership	30	0.02
491	professional	30	0.02
492	saw	30	0.02
493	sec	30	0.02
494	want	30	0.02
495	yard	30	0.02
496	Andrew	29	0.02
497	around	29	0.02
498	Auckland	29	0.02
499	Baltimore	29	0.02
500	central	29	0.02
501	decision	29	0.02
502	Diego	29	0.02
503	Eagles	29	0.02
504	every	29	0.02
505	few	29	0.02
506	keep	29	0.02
507	Scott	29	0.02
508	several	29	0.02
509	streak	29	0.02
510	taking	29	0.02
511	TV	29	0.02
512	able	28	0.02
513	bowling	28	0.02
514	Def	28	0.02
515	Eastern	28	0.02
516	events	28	0.02
517	James	28	0.02
518	late	28	0.02
519	nations	28	0.02
520	NBA	28	0.02
521	success	28	0.02
522	training	28	0.02
523	twice	28	0.02
524	winners	28	0.02

Rank	Word	Freq.	Percent
525	Wayne	28	0.02
526	Brazilian	27	0.02
527	Brisbane	27	0.02
528	Ferrari	27	0.02
529	formula	27	0.02
530	free	27	0.02
531	Netherlands	27	0.02
532	October	27	0.02
533	Olympic	27	0.02
534	period	27	0.02
535	Phoenix	27	0.02
536	picked	27	0.02
537	pressure	27	0.02
538	serve	27	0.02
539	Slovakia	27	0.02
540	squad	27	0.02
541	trophy	27	0.02
542	Venus	27	0.02
543	Washington	27	0.02
544	announced	26	0.02
545	become	26	0.02
546	forward	26	0.02
547	Knicks	26	0.02
548	Mike	26	0.02
549	named	26	0.02
550	racing	26	0.02
551	stars	26	0.02
552	started	26	0.02
553	sure	26	0.02
554	bit	25	0.02
555	car	25	0.02
556	full	25	0.02
557	given	25	0.02
558	Jason	25	0.02
559	Jose	25	0.02
560	knee	25	0.02
561	level	25	0.02
562	Lewis	25	0.02
563	may	25	0.02
564	medal	25	0.02
565	missed	25	0.02
566	move	25	0.02
567	Portugal	25	0.02
568	Rus	25	0.02
569	Russian	25	0.02
570	saves	25	0.02
571	Schumacher	25	0.02
572	Scotland	25	0.02
573	Serena	25	0.02
574	shots	25	0.02
575	striker	25	0.02
576	such	25	0.02
577	touchdown	25	0.02
578	Venables	25	0.02
579	whose	25	0.02
580	already	24	0.02
581	Andre	24	0.02
582	assists	24	0.02
583	award	24	0.02

Rank	Word	Freq.	Percent
584	Davis	24	0.02
585	far	24	0.02
586	Fra	24	0.02
587	Giants	24	0.02
588	history	24	0.02
589	hole	24	0.02
590	looked	24	0.02
591	Malaysia	24	0.02
592	Ohio	24	0.02
593	overall	24	0.02
594	Peter	24	0.02
595	probably	24	0.02
596	return	24	0.02
597	Safin	24	0.02
598	seventh	24	0.02
599	Slam	24	0.02
600	Supersport	24	0.02
601	Swiss	24	0.02
602	tough	24	0.02
603	Wales	24	0.02
604	Alex	23	0.02
605	beaten	23	0.02
606	became	23	0.02
607	channel	23	0.02
608	delayed	23	0.02
609	dropped	23	0.02
610	earned	23	0.02
611	eighth	23	0.02
612	head	23	0.02
613	helped	23	0.02
614	hope	23	0.02
615	Ireland	23	0.02
616	Jan	23	0.02
617	join	23	0.02
618	Lee	23	0.02
619	line	23	0.02
620	medals	23	0.02
621	min	23	0.02
622	nothing	23	0.02
623	previous	23	0.02
624	row	23	0.02
625	Rusedski	23	0.02
626	says	23	0.02
627	Spanish	23	0.02
628	tied	23	0.02
629	titles	23	0.02
630	beating	22	0.02
631	believe	22	0.02
632	Blacks	22	0.02
633	Busan	22	0.02
634	centre	22	0.02
635	college	22	0.02
636	contract	22	0.02
637	due	22	0.02
638	enough	22	0.02
639	failed	22	0.02
640	fell	22	0.02
641	German	22	0.02
642	hopes	22	0.02

Rank	Word	Freq.	Percent
643	injured	22	0.02
644	Italian	22	0.02
645	km	22	0.02
646	love	22	0.02
647	main	22	0.02
648	miss	22	0.02
649	Oklahoma	22	0.02
650	power	22	0.02
651	Rangers	22	0.02
652	running	22	0.02
653	say	22	0.02
654	super	22	0.02
655	those	22	0.02
656	Tiger	22	0.02
657	Tim	22	0.02
658	veteran	22	0.02
659	Walker	22	0.02
660	America	21	0.02
661	broke	21	0.02
662	chances	21	0.02
663	Chinese	21	0.02
664	conference	21	0.02
665	consecutive	21	0.02
666	cycling	21	0.02
667	difficult	21	0.02
668	does	21	0.02
669	draw	21	0.02
670	easy	21	0.02
671	felt	21	0.02
672	Ferrero	21	0.02
673	Finland	21	0.02
674	front	21	0.02
675	Georgia	21	0.02
676	Green	21	0.02
677	Jersey	21	0.02
678	kept	21	0.02
679	less	21	0.02
680	let	21	0.02
681	morning	21	0.02
682	rivals	21	0.02
683	Smith	21	0.02
684	something	21	0.02
685	starting	21	0.02
686	swimming	21	0.02
687	taken	21	0.02
688	unbeaten	21	0.02
689	why	21	0.02
690	yet	21	0.02
691	among	20	0.02
692	Antonio	20	0.02
693	based	20	0.02
694	bowl	20	0.02
695	boxing	20	0.02
696	Chi	20	0.02
697	circuit	20	0.02
698	countries	20	0.02
699	current	20	0.02
700	defeated	20	0.02
701	drive	20	0.02

Rank	Word	Freq.	Percent
702	driver	20	0.02
703	Edwards	20	0.02
704	Fernando	20	0.02
705	finally	20	0.02
706	foot	20	0.02
707	Ford	20	0.02
708	Gbr	20	0.02
709	ground	20	0.02
710	overtime	20	0.02
711	Paris	20	0.02
712	pool	20	0.02
713	pulled	20	0.02
714	September	20	0.02
715	single	20	0.02
716	stay	20	0.02
717	suffered	20	0.02
718	surgery	20	0.02
719	track	20	0.02
720	trying	20	0.02
721	Ukraine	20	0.02
722	until	20	0.02
723	upset	20	0.02
724	used	20	0.02
725	victories	20	0.02
726	Wayne	20	0.02
727	WTA	20	0.02
728	young	20	0.02
729	your	20	0.02
730	Afridi	19	0.01
731	baht	19	0.01
732	battle	19	0.01
733	began	19	0.01
734	bottom	19	0.01
735	California	19	0.01
736	center	19	0.01
737	conditions	19	0.01
738	crown	19	0.01
739	Daniel	19	0.01
740	expected	19	0.01
741	fastest	19	0.01
742	Ferguson	19	0.01
743	Ger	19	0.01
744	golden	19	0.01
745	happy	19	0.01
746	Henman	19	0.01
747	himself	19	0.01
748	Hurricanes	19	0.01
749	interest	19	0.01
750	Iowa	19	0.01
751	Lakers	19	0.01
752	Liverpool	19	0.01
753	Michigan	19	0.01
754	moved	19	0.01
755	Munich	19	0.01
756	need	19	0.01
757	OneWorld	19	0.01
758	opened	19	0.01
759	police	19	0.01
760	proved	19	0.01

Rank	Word	Freq.	Percent
761	quarterback	19	0.01
762	rankings	19	0.01
763	ready	19	0.01
764	rest	19	0.01
765	short	19	0.01
766	successful	19	0.01
767	asked	18	0.01
768	Austria	18	0.01
769	bad	18	0.01
770	Bangladesh	18	0.01
771	birdies	18	0.01
772	Caps	18	0.01
773	com	18	0.01
774	completed	18	0.01
775	confident	18	0.01
776	debut	18	0.01
777	December	18	0.01
778	different	18	0.01
779	ended	18	0.01
780	extras	18	0.01
781	freestyle	18	0.01
782	Gonzalez	18	0.01
783	Hungary	18	0.01
784	improved	18	0.01
785	individual	18	0.01
786	injuries	18	0.01
787	Kevin	18	0.01
788	lose	18	0.01
789	meeting	18	0.01
790	Miller	18	0.01
791	official	18	0.01
792	opener	18	0.01
793	pair	18	0.01
794	percent	18	0.01
795	Raptors	18	0.01
796	rookie	18	0.01
797	seen	18	0.01
798	sets	18	0.01
799	Singapore	18	0.01
800	special	18	0.01
801	statement	18	0.01
802	step	18	0.01
803	Tennessee	18	0.01
804	thanks	18	0.01
805	Warriors	18	0.01
806	Webber	18	0.01
807	worst	18	0.01
808	Adelaide	17	0.01
809	admitted	17	0.01
810	along	17	0.01
811	athletics	17	0.01
812	August	17	0.01
813	basketball	17	0.01
814	biggest	17	0.01
815	Brian	17	0.01
816	charges	17	0.01
817	Colin	17	0.01
818	Coughlin	17	0.01
819	decided	17	0.01

Rank	Word	Freq.	Percent
820	Ed	17	0.01
821	Faldo	17	0.01
822	fine	17	0.01
823	force	17	0.01
824	Hantuchova	17	0.01
825	important	17	0.01
826	Japanese	17	0.01
827	King	17	0.01
828	leader	17	0.01
829	least	17	0.01
830	life	17	0.01
831	McNabb	17	0.01
832	might	17	0.01
833	pace	17	0.01
834	Pollock	17	0.01
835	prize	17	0.01
836	rally	17	0.01
837	ranking	17	0.01
838	road	17	0.01
839	scheduled	17	0.01
840	successive	17	0.01
841	Taylor	17	0.01
842	tournaments	17	0.01
843	turn	17	0.01
844	Younis	17	0.01
845	according	16	0.01
846	aggregate	16	0.01
847	air	16	0.01
848	Alinghi	16	0.01
849	allowed	16	0.01
850	Angels	16	0.01
851	ankle	16	0.01
852	Armstrong	16	0.01
853	blue	16	0.01
854	board	16	0.01
855	body	16	0.01
856	Brad	16	0.01
857	bring	16	0.01
858	broken	16	0.01
859	Byrd	16	0.01
860	called	16	0.01
861	campaign	16	0.01
862	class	16	0.01
863	classic	16	0.01
864	closer	16	0.01
865	Croatia	16	0.01
866	darts	16	0.01
867	deal	16	0.01
868	Denver	16	0.01
869	dream	16	0.01
870	edged	16	0.01
871	faces	16	0.01
872	fact	16	0.01
873	Fin	16	0.01
874	fired	16	0.01
875	Gillespie	16	0.01
876	homer	16	0.01
877	Honda	16	0.01
878	Hussain	16	0.01

Rank	Word	Freq.	Percent
879	Ivan	16	0.01
880	Jamaica	16	0.01
881	job	16	0.01
882	Jordan	16	0.01
883	Kent	16	0.01
884	Kenya	16	0.01
885	known	16	0.01
886	Korean	16	0.01
887	laps	16	0.01
888	metre	16	0.01
889	Moya	16	0.01
890	Newcastle	16	0.01
891	Nicolas	16	0.01
892	November	16	0.01
893	opponent	16	0.01
894	perfect	16	0.01
895	position	16	0.01
896	produced	16	0.01
897	rain	16	0.01
898	regular	16	0.01
899	returned	16	0.01
900	Roy	16	0.01
901	runner	16	0.01
902	save	16	0.01
903	saying	16	0.01
904	showed	16	0.01
905	Swe	16	0.01
906	Taiwan	16	0.01
907	takes	16	0.01
908	Tech	16	0.01
909	these	16	0.01
910	things	16	0.01
911	Thomas	16	0.01
912	Thorpe	16	0.01
913	Yao	16	0.01
914	Zhang	16	0.01
915	African	15	0.01
916	almost	15	0.01
917	Americans	15	0.01
918	Andy	15	0.01
919	Ashes	15	0.01
920	balls	15	0.01
921	Belgian	15	0.01
922	believes	15	0.01
923	Ben	15	0.01
924	Bulldogs	15	0.01
925	call	15	0.01
926	Canadian	15	0.01
927	caught	15	0.01
928	change	15	0.01
929	charged	15	0.01
930	clubs	15	0.01
931	Dallas	15	0.01
932	defence	15	0.01
933	done	15	0.01
934	East	15	0.01
935	either	15	0.01
936	Eric	15	0.01
937	everyone	15	0.01

Rank	Word	Freq.	Percent
938	experience	15	0.01
939	Federation	15	0.01
940	finishing	15	0.01
941	fought	15	0.01
942	found	15	0.01
943	general	15	0.01
944	George	15	0.01
945	hold	15	0.01
946	host	15	0.01
947	Indiana	15	0.01
948	Ita	15	0.01
949	jump	15	0.01
950	junior	15	0.01
951	Kafelnikov	15	0.01
952	Keane	15	0.01
953	knew	15	0.01
954	LA	15	0.01
955	lack	15	0.01
956	Meara	15	0.01
957	media	15	0.01
958	midfielder	15	0.01
959	mind	15	0.01
960	months	15	0.01
961	Oliver	15	0.01
962	organisers	15	0.01
963	Poland	15	0.01
964	post	15	0.01
965	Premier	15	0.01
966	problems	15	0.01
967	putting	15	0.01
968	qualifying	15	0.01
969	reason	15	0.01
970	received	15	0.01
971	referee	15	0.01
972	reported	15	0.01
973	rival	15	0.01
974	Romania	15	0.01
975	seasons	15	0.01
976	sent	15	0.01
977	signing	15	0.01
978	skating	15	0.01
979	stroke	15	0.01
980	surprise	15	0.01
981	thing	15	0.01
982	thought	15	0.01
983	Todd	15	0.01
984	turned	15	0.01
985	Tyson	15	0.01
986	Yugoslavia	15	0.01
987	anything	14	0.01
988	attack	14	0.01
989	Aus	14	0.01
990	birdie	14	0.01
991	card	14	0.01
992	Cardinals	14	0.01
993	Carolina	14	0.01
994	certainly	14	0.01
995	Chittinand	14	0.01
996	Christian	14	0.01

Rank	Word	Freq.	Percent
997	claimed	14	0.01
998	cost	14	0.01
999	Coyotes	14	0.01
1000	denotes	14	0.01
1001	Edmonton	14	0.01
1002	Europe	14	0.01
1003	favourite	14	0.01
1004	figure	14	0.01
1005	followed	14	0.01
1006	further	14	0.01
1007	Greece	14	0.01
1008	hand	14	0.01
1009	holding	14	0.01
1010	horse	14	0.01
1011	hour	14	0.01
1012	impressive	14	0.01
1013	Indianapolis	14	0.01
1014	joined	14	0.01
1015	June	14	0.01
1016	kick	14	0.01
1017	Kings	14	0.01
1018	Kournikova	14	0.01
1019	Lance	14	0.01
1020	live	14	0.01
1021	Lleyton	14	0.01
1022	look	14	0.01
1023	low	14	0.01
1024	managed	14	0.01
1025	maybe	14	0.01
1026	mixed	14	0.01
1027	moment	14	0.01
1028	myself	14	0.01
1029	Nets	14	0.01
1030	news	14	0.01
1031	Olympics	14	0.01
1032	order	14	0.01
1033	pitch	14	0.01
1034	playoffs	14	0.01
1035	Pol	14	0.01
1036	problem	14	0.01
1037	qualifier	14	0.01
1038	rather	14	0.01
1039	reached	14	0.01
1040	recently	14	0.01
1041	reporters	14	0.01
1042	retired	14	0.01
1043	Robert	14	0.01
1044	Robinson	14	0.01
1045	Rockets	14	0.01
1046	runners	14	0.01
1047	security	14	0.01
1048	Shanghai	14	0.01
1049	show	14	0.01
1050	Simon	14	0.01
1051	skipper	14	0.01
1052	Snow	14	0.01
1053	starts	14	0.01
1054	swimmers	14	0.01
1055	table	14	0.01

Rank	Word	Freq.	Percent
1056	Thais	14	0.01
1057	Thongchai	14	0.01
1058	throughout	14	0.01
1059	Tor	14	0.01
1060	touchdowns	14	0.01
1061	Tpe	14	0.01
1062	transfer	14	0.01
1063	true	14	0.01
1064	Union	14	0.01
1065	use	14	0.01
1066	wind	14	0.01
1067	agent	13	0.01
1068	agreed	13	0.01
1069	appeared	13	0.01
1070	Arsenal	13	0.01
1071	backstroke	13	0.01
1072	ban	13	0.01
1073	bid	13	0.01
1074	Blake	13	0.01
1075	Bortolo	13	0.01
1076	boss	13	0.01
1077	Boucher	13	0.01
1078	boxers	13	0.01
1079	breaking	13	0.01
1080	brought	13	0.01
1081	camp	13	0.01
1082	challengers	13	0.01
1083	clear	13	0.01
1084	Collins	13	0.01
1085	Commonwealth	13	0.01
1086	confidence	13	0.01
1087	daily	13	0.01
1088	Darren	13	0.01
1089	dominated	13	0.01
1090	Donald	13	0.01
1091	especially	13	0.01
1092	fan	13	0.01
1093	Fassa	13	0.01
1094	fast	13	0.01
1095	Federer	13	0.01
1096	fellow	13	0.01
1097	friendly	13	0.01
1098	Gen	13	0.01
1099	Gibbs	13	0.01
1100	Gilchrist	13	0.01
1101	grabbed	13	0.01
1102	Indies	13	0.01
1103	Jelena	13	0.01
1104	July	13	0.01
1105	Kelme	13	0.01
1106	latest	13	0.01
1107	leaving	13	0.01
1108	likely	13	0.01
1109	local	13	0.01
1110	Martinez	13	0.01
1111	Matta	13	0.01
1112	McRae	13	0.01
1113	missing	13	0.01
1114	NHL	13	0.01

Rank	Word	Freq.	Percent
1115	Nick	13	0.01
1116	often	13	0.01
1117	Oracle	13	0.01
1118	pairs	13	0.01
1119	pass	13	0.01
1120	Patrick	13	0.01
1121	per	13	0.01
1122	playoff	13	0.01
1123	pretty	13	0.01
1124	quickly	13	0.01
1125	quite	13	0.01
1126	Rios	13	0.01
1127	Rose	13	0.01
1128	rules	13	0.01
1129	seeding	13	0.01
1130	seems	13	0.01
1131	selected	13	0.01
1132	Sharks	13	0.01
1133	Spurs	13	0.01
1134	Srichaphan	13	0.01
1135	Terry	13	0.01
1136	trial	13	0.01
1137	trouble	13	0.01
1138	try	13	0.01
1139	unless	13	0.01
1140	wanted	13	0.01
1141	whole	13	0.01
1142	within	13	0.01

APPENDIX B

Word Frequency List in Each Category of Sports News

AS = Advance Stories
 SAS = Sports Activity Schedules
 IA = Interviewing Articles
 FES = Following Event Stories
 JC = Journalist Comments
 SaS = Scores and Standings
 RO = Reader Opinions

Rank	Overall	AS	FES	SaS	SAS	JC	RO	IA
	Word	Word	Word	Word	Word	Word	Word	Word
	Freq.	Freq.	Freq.	Freq.	Freq.	Freq.	Freq.	Freq.
1	the	1,568	4,049	190	V	286	the	329
2	in	640	1,617	120	ESPN	115	to	134
3	to	613	1,511	117	golf	109	and	110
4	a	2,346	1,380	107	star	108	of	92
5	and	2,124	1,237	99	Supersport	104	A	76
6	of	2,012	1,150	98	day	99	in	75
7	's	1,415	726	92	delayed	64	it	68
8	for	1,112	665	92	soccer	60	is	54
9	on	1,039	580	92	a.m.	57	that	54
10	was	753	518	88	Tour	50	's	53
11	with	738	476	88	PGA	47	I	51
12	at	722	426	87	baseball	45	for	50
13	his	637	385	86	championship	43	be	40
14	I	626	354	68	channel	37	have	39
15	he	600	352	66	League	37	on	39
16	that	545	324	65	major	36	as	37
17	it	534	307	57	on	31	world	37
18	as	523	279	56	TV	30	football	32
19	is	521	265	55	p.m.	30	you	31
20	said	508	264	51	today	28	not	30

Rank	Overall		AS.txt		FES.txt		SaS.txt		SAS.txt		JC.txt		RO.txt		IA.txt	
	Word	Freq.	Word	Freq.	Word	Freq.	Word	Freq.	Word	Freq.	Word	Freq.	Word	Freq.	Word	Freq.
21	first	476	year	151	but	260	men	51	Open	13	have	28	n't	29	from	19
22	from	467	was	146	after	254	metres	51	tennis	12	not	28	this	29	Thailand	19
23	but	438	it	144	from	240	women	50	Thailand	12	their	28	they	28	have	18
24	by	415	have	143	win	240	bt	46	ATP	9	Thai	27	we	27	Paradorn	17
25	be	413	from	123	by	239	NY	45	European	8	but	25	game	26	world	17
26	have	411	as	119	two	239	from	44	morning	8	games	25	with	26	said	16
27	year	404	but	119	who	239	standings	43	series	8	this	25	are	25	with	16
28	will	394	who	107	second	230	Russia	42	Masters	7	which	25	by	25	but	14
29	who	390	team	101	over	229	Spa	41	world	7	who	24	what	24	skating	14
30	had	378	this	101	that	228	overs	40	cycling	6	his	23	would	23	they	14
31	two	365	not	98	year	222	matches	39	France	6	on	23	cup	22	this	14
32	after	364	last	96	No.	212	out	39	Monday	6	when	23	if	22	are	13
33	world	364	world	94	points	207	St.	38	Thursday	6	by	22	sports	22	my	13
34	one	360	one	90	one	203	Canada	37	Cardinals	5	about	19	but	20	year	13
35	has	346	by	86	game	197	round	37	championships	5	all	19	so	20	Asia	12
36	not	327	are	85	world	189	points	36	Cincinnati	5	athletes	19	he	17	by	12
37	their	307	we	85	their	184	first	35	Cup	5	has	19	more	17	him	12
38	win	300	season	83	have	183	south	35	English	5	some	19	their	17	Nathapong	11
39	this	297	been	80	three	175	Czech	34	football	5	Thailand	19	was	17	who	11
40	beat	294	two	79	out	171	one	34	Friday	5	an	18	can	16	Thai	10
41	out	279	after	77	last	169	Republic	34	game	5	so	18	will	16	up	10
42	over	279	had	77	is	167	minutes	32	ITV	5	Asian	17	all	15	all	9
43	second	276	first	74	season	161	tennis	32	Manchester	5	officials	17	at	15	at	9
44	No.	274	would	74	won	160	by	31	Premiership	5	there	17	ed	15	Cup	9
45	an	269	their	71	an	159	Michael	31	St.	5	were	17	team	15	match	9
46	last	269	they	68	time	155	national	31	Tournament	5	will	17	there	15	so	9
47	game	268	up	67	up	146	Saturday	31	united	5	Venables	16	about	14	n't	9
48	team	264	an	65	final	142	Africa	30	and	4	only	15	been	14	there	9
49	season	259	which	65	victory	141	USA	30	Asean	4	sport	15	his	14	we	9
50	they	259	Australia	63	four	140	time	29	Buick	4	sports	15	from	13	against	8

Rank	Overall		AS.txt		FES.txt		SaS.txt		SAS.txt		JC.txt		RO.txt		IA.txt	
	Word	Freq.	Word	Freq.	Word	Freq.	Word	Freq.	Word	Freq.	Word	Freq.	Word	Freq.	Word	Freq.
51	points	258	US	63	not	140	Chicago	28	final	4	can	14	much	13	Agassi	8
52	final	252	players	61	round	140	def	28	international	4	I	14	one	13	an	8
53	are	249	Cup	60	be	138	Sunday	28	NFL	4	like	14	players	13	been	8
54	time	249	all	59	has	135	Belgium	27	pool	4	one	14	two	13	being	8
55	we	245	England	59	when	133	Cincinnati	27	professional	4	well	14	how	12	made	8
56	all	243	Thailand	59	US	132	David	27	sports	4	you	14	should	12	professional	8
57	three	242	against	58	her	131	not	27	this	4	been	13	Thai	12	she	8
58	up	237	yesterday	57	match	131	western	27	to	4	year	13	them	12	skate	8
59	when	230	next	56	scored	128	yesterday	27	youth	4	n't	12	after	11	team	8
60	Australia	228	when	56	third	123	American	25	American	3	just	11	Fifa	11	well	8
61	US	226	three	52	this	123	Switzerland	25	Arizona	3	many	11	has	11	which	8
62	won	224	Reuters	50	five	122	Toronto	25	Arsenal	3	or	11	well	11	win	8
63	round	208	out	46	were	121	Argentina	24	Centenary	3	training	11	when	11	because	7
64	England	207	play	46	all	120	Arizona	24	Diamondbacks	3	Association	10	an	10	center	7
65	Cup	205	also	45	we	116	Florida	24	division	3	could	10	Bangkok	10	come	7
66	four	203	before	45	just	115	Mets	24	first	3	countries	10	like	10	football	7
67	match	203	time	45	team	115	Milwaukee	24	Francisco	3	even	10	only	10	ice	7
68	n't	203	match	44	they	112	Montreal	24	friendly	3	if	10	or	10	masters	7
69	league	201	Asian	43	against	111	played	24	German	3	other	10	real	10	old	7
70	yesterday	197	club	43	yesterday	111	result	24	Giants	3	out	10	were	10	rink	7
71	which	190	million	43	Reuters	110	White	24	Liverpool	3	time	10	any	9	serve	7
72	m	189	champion	42	will	108	Atlanta	23	Madrid	3	world	10	final	9	when	7
73	games	184	No.	42	lead	107	Boston	23	motor	3	had	9	good	9	WISC	7
74	been	183	old	42	off	107	city	23	racing	3	medals	9	just	9	women	7
75	Thailand	182	there	42	n't	105	Colorado	23	Scottish	3	now	9	out	9	would	7
76	against	181	top	42	set	102	Cubs	23	under	3	she	9	people	9	after	6
77	were	180	about	41	back	101	eastern	23	US	3	two	9	see	9	also	6
78	just	167	four	41	run	101	Houston	23	Wimbledon	3	amateur	8	Sportsgag	9	can	6
79	Open	166	grand	41	while	101	Minnesota	23	a	2	being	8	way	9	commercial	6
80	third	162	League	41	games	98	Philadelphia	23	amateur	2	better	8	who	9	first	6

Rank	Overall		AS.txt		FES.txt		SaS.txt		SAS.txt		JC.txt		RO.txt		IA.txt	
	Word	Freq.	Word	Freq.	Word	Freq.	Word	Freq.	Word	Freq.	Word	Freq.	Word	Freq.	Word	Freq.
81	victory	162	n't	41	also	97	Pittsburgh	23	Amsterdam	2	both	8	years	9	get	6
82	also	160	final	39	into	96	Rus	23	Atlanta	2	boxers	8	did	8	into	6
83	Reuters	160	Open	39	top	95	Sweden	23	Borussia	2	come	8	do	8	me	6
84	would	156	years	38	Cup	94	Baltimore	22	Braves	2	do	8	does	8	one	6
85	her	154	him	37	old	92	baseball	22	British	2	Gen	8	goals	8	player	6
86	top	152	off	37	title	89	central	22	change	2	Henman	8	new	8	take	6
87	day	150	back	36	beat	87	day	22	circuit	2	most	8	No.	8	their	6
88	off	150	championship	36	she	86	Fra	22	city	2	No.	8	play	8	while	6
89	five	149	games	36	only	85	group	22	Classic	2	Olympics	8	teams	8	amateur	5
90	old	149	more	36	him	84	Kansas	22	Cleveland	2	players	8	Thailand	8	both	5
91	south	148	former	35	record	84	Netherlands	22	Community	2	said	8	than	8	country	5
92	only	147	Tournament	34	which	84	Oakland	22	Dodgers	2	up	8	which	8	court	5
93	play	146	won	34	goal	83	Seattle	22	events	2	we	8	win	8	even	5
94	back	144	only	33	Open	83	team	22	free	2	after	7	had	7	event	5
95	Tournament	142	over	33	play	82	Anaheim	21	guide	2	England	7	Keane	7	getting	5
96	him	141	week	33	are	81	ATP	21	Hungarian	2	gold	7	let	7	had	5
97	while	141	player	32	added	80	bay	21	Indians	2	international	7	make	7	hard	5
98	players	140	second	32	home	79	Detroit	21	Indonesia	2	more	7	playing	7	having	5
99	there	138	Tour	32	before	78	sec	21	Leeds	2	much	7	some	7	just	5
100	before	135	international	31	fourth	78	Texas	21	match	2	must	7	south	7	Kingdom	5

APPENDIX C

Content Words in Sports News Corpus

1. Noun

No.	Noun	Freq.	Percent
1	year	404	3.09
2	world	364	2.78
3	game	268	2.05
4	team	262	2.01
5	season	259	1.98
6	points	258	1.97
7	time	249	1.90
8	Australia	228	1.75
9	round	208	1.59
10	England	207	1.58
11	Cup	205	1.57
12	League	201	1.54
13	match	199	1.52
14	final	198	1.52
15	games	184	1.41
16	Thailand	182	1.39
17	win	177	1.35
18	Open	163	1.25
19	victory	162	1.24
20	Reuters	160	1.22
21	day	150	1.15
22	Tournament	142	1.09
23	players	140	1.07
24	division	131	1.00
25	home	128	0.98
26	Spain	126	0.96
27	champion	124	0.95
28	France	124	0.95
29	title	121	0.93
30	results	120	0.92
31	Tour	118	0.90
32	women	117	0.89
33	seconds	113	0.86
34	player	110	0.84
35	championship	109	0.83
36	record	107	0.82
37	minutes	101	0.77
38	run	97	0.74
39	Saturday	97	0.74
40	finals	96	0.73
41	men	95	0.73
42	Germany	94	0.72
43	football	92	0.70
44	years	92	0.70
45	goal	91	0.70
46	place	90	0.69
47	week	89	0.68
48	golf	88	0.67
49	quarter	87	0.67
50	Sunday	87	0.67

No.	Noun	Freq.	Percent
51	million	86	0.66
52	set	84	0.64
53	matches	83	0.63
54	Paradorn	83	0.63
55	lead	81	0.62
56	Italy	81	0.62
57	club	78	0.60
58	Michael	78	0.60
59	teams	78	0.60
60	tennis	77	0.59
61	metres	76	0.58
62	star	75	0.57
63	race	74	0.57
64	series	74	0.57
65	Williams	74	0.57
66	number	73	0.56
67	side	73	0.56
68	Africa	72	0.55
69	champions	72	0.55
70	seed	70	0.54
71	event	68	0.52
72	Japan	68	0.52
73	coach	67	0.51
74	runs	66	0.50
75	overs	65	0.50
76	career	63	0.48
77	goals	63	0.48
78	sports	63	0.48
79	way	63	0.48
80	innings	62	0.47
81	London	62	0.47
82	Friday	61	0.47
83	ball	59	0.45
84	city	59	0.45
85	Mark	59	0.45
86	Tuesday	59	0.45
87	David	58	0.44
88	Russia	58	0.44
89	winner	58	0.44
90	injury	57	0.44
91	Monday	57	0.44
92	standings	57	0.44
93	Toronto	57	0.44
94	par	56	0.43
95	Woods	56	0.43
96	Chicago	55	0.42
97	fans	55	0.42
98	field	55	0.42
99	month	55	0.42
100	yards	55	0.42
101	championships	54	0.41
102	chance	54	0.41
103	China	54	0.41
104	Florida	54	0.41
105	people	54	0.41
106	Brazil	53	0.41
107	Canada	51	0.39
108	group	51	0.39
109	Texas	51	0.39

No.	Noun	Freq.	Percent
110	United	51	0.39
111	baseball	50	0.38
112	half	50	0.38
113	Boston	50	0.38
114	point	50	0.38
115	rebounds	50	0.38
116	soccer	50	0.38
117	Masters	49	0.37
118	Philadelphia	49	0.37
119	Cincinnati	48	0.37
120	end	48	0.37
121	Korea	48	0.37
122	man	48	0.37
123	sport	48	0.37
124	Atlanta	47	0.36
125	course	47	0.36
126	manager	47	0.36
127	shot	47	0.36
128	start	47	0.36
129	Wednesday	47	0.36
130	Arizona	46	0.35
131	Britain	46	0.35
132	minute	46	0.35
133	Montreal	46	0.35
134	states	46	0.35
135	test	46	0.35
136	times	46	0.35
137	play	38	0.29
138	defeat	36	0.28
139	gold	21	0.16
140	best	20	0.15
141	today	14	0.11
142	hit	10	0.08
143	playing	8	0.06
144	winning	8	0.06
145	going	2	0.02
Total		13,078	100

2. Verb

No.	Verb	Freq.	Percent
1	said	508	15.53
2	beat	294	8.99
3	won	224	6.85
4	scored	133	4.07
5	win	123	3.76
6	played	117	3.58
7	play	108	3.30
8	made	107	3.27
9	added	94	2.87
10	lost	85	2.60
11	hit	79	2.42
12	get	78	2.38
13	take	78	2.38
14	do	75	2.29
15	got	72	2.20
16	going	69	2.11
17	took	66	2.02
18	come	65	1.99
19	go	65	1.99
20	playing	64	1.96
21	did	61	1.86
22	make	60	1.83
23	think	59	1.80
24	put	54	1.65
25	winning	50	1.53
26	came	49	1.50
27	finished	49	1.50
28	went	49	1.50
29	told	48	1.47
30	ranked	47	1.44
31	set	47	1.44
32	left	42	1.28
33	lead	37	1.13
34	run	21	0.64
35	shot	17	0.52
36	leading	16	0.49
37	end	13	0.40
38	defeat	12	0.37
39	start	10	0.31
40	coach	4	0.12
41	match	4	0.12
42	opening	4	0.12
43	open	3	0.09
44	race	3	0.09
45	record	3	0.09
46	team	2	0.06
47	test	2	0.06
48	place	1	0.03
Total		3,271	100

3. Adjective

No.	Adjective	Freq.	Percent
1	new	267	11.06
2	top	152	6.30
3	old	149	6.17
4	south	148	6.13
5	Thai	129	5.34
6	next	118	4.89
7	American	111	4.59
8	national	97	4.02
9	good	90	3.73
10	Asian	89	3.69
11	grand	85	3.52
12	international	78	3.23
13	former	74	3.07
14	best	71	2.94
15	Australian	67	2.78
16	Major	67	2.78
17	big	57	2.36
18	English	55	2.28
19	long	55	2.28
20	final	54	2.24
21	European	53	2.19
22	opening	53	2.19
23	United	51	2.11
24	French	48	1.99
25	winning	41	1.69
26	early	34	1.41
27	gold	34	1.41
28	leading	33	1.37
29	left	26	1.08
30	half	23	0.95
31	earlier	5	0.21
Total		2,414	100

4. Adverb

No.	Adverb	Freq.	Percent
1	yesterday	197	14.05
2	just	167	11.91
3	also	160	11.41
4	only	147	10.49
5	now	100	7.13
6	well	94	6.71
7	here	87	6.21
8	still	69	4.92
9	then	65	4.64
10	today	63	4.49
11	again	58	4.14
12	away	50	3.57
13	really	42	2.99
14	never	41	2.92
15	earlier	36	2.57
16	early	26	1.85
Total		1,402	100

APPENDIX D

Abbreviations and Acronyms in Sports News Corpus

Rank	Word	Full Name	Freq.	Percent
1	's	is, has	1,415	34.12
2	n't	not	327	7.89
3	No.	Number	274	6.61
4	US	United States	226	5.45
5	St.	Saint, State	115	2.77
6	AFP	Associated French Press	80	1.93
7	m	metre	67	1.62
8	a.m.	ante meridiem	66	1.59
9	'm	am	57	1.37
10	've	have	57	1.37
11	bt	beat	51	1.23
12	Spa	Spain	49	1.18
13	ATP	Association of Tennis Professionals	46	1.11
14	NY	New York	46	1.11
15	V	Versus	46	1.11
16	PGA	Professional Golf Association	45	1.09
17	USA	United States of America	38	0.92
18	ESPN	Entertainment and Sports Programming Network	36	0.87
19	AP	Associated Press	35	0.84
20	p.m.	post meridiem	30	0.72
21	sec	second	30	0.72
22	TV	television	29	0.70
23	def	defeated	28	0.68
24	NBA	National Basketball Association	28	0.68
25	Rus	Russia	25	0.60
26	Fra	France	24	0.58
27	min	minute	23	0.55
28	km	kilometre	22	0.53
29	Chi	Chicago	20	0.48
30	GBR	Great Britain	20	0.48
31	WTA	Women's Tennis Association	20	0.48
32	Ger	Germany	19	0.46
33	caps	capitals	18	0.43
34	Ed	Editor	17	0.41
35	Fin	Finland	16	0.39
36	Swe	Sweden	16	0.39
37	Ita	Italy	15	0.36
38	LA	Los Angeles	15	0.36
39	Aus	Australia	14	0.34
40	Pol	Poland	14	0.34
41	Tpe	Taipei	14	0.34
42	FI	Formula One	13	0.31
43	Gen	General	13	0.31
44	NHL	National Hockey League	13	0.31
45	FIFA	International Association of Football Federations	12	0.29
46	hr	hour	12	0.29
47	IBF	International Badminton Federation (6) International Boxing Federation (6)	12	0.29
48	RSA	Republic of South Africa	12	0.29
49	Bra	Brazil	11	0.27
50	NFL	National Football League	11	0.27
51	Eng	England	10	0.24
52	Ind	India	10	0.24

Rank	Word	Full Name	Freq.	Percent
53	LPGA	Lady Player Golfers Association	10	0.24
54	Mr.	Mister	10	0.24
55	Tha	Thailand	10	0.24
56	Arg	Argentina	9	0.22
57	Cze	Czech Republic	9	0.22
58	NZ	New Zealand	9	0.22
59	OT	overtime	9	0.22
60	secs	seconds	9	0.22
61	TD	touchdown	9	0.22
62	CART	Championship Auto Racing Teams	8	0.19
63	FA	Football Association	8	0.19
64	Kor	Korea	8	0.19
65	mph	mile per hour	8	0.19
66	Ned	Netherlands	8	0.19
67	pts	points	8	0.19
68	rd	round	8	0.19
69	TALGA	Thailand Amateur Ladies Golf Association	7	0.17
70	WISC	World Ice Skating Center	7	0.17
71	Aug	August	6	0.14
72	Chn	China	6	0.14
73	Gre	Greece	6	0.14
74	IAAF	International Association of Athletics Federations	6	0.14
75	Jpn	Japan	6	0.14
76	Lit	Lithuania	6	0.14
77	Pro	Professional	6	0.14
78	Sept	September	6	0.14
79	ABAT	Amateur Boxing Association of Thailand	5	0.12
80	ARFU	Asian Rugby Football Union	5	0.12
81	CC	Cubic Centimeter	5	0.12
82	IRL	Indy Racing League	5	0.12
83	NL	National League	5	0.12
84	Nov	November	5	0.12
85	QB	quarterback	5	0.12
86	ASEAN	Association of South-East Asian Nations	4	0.10
87	BBC	British Broadcasting Corporation	4	0.10
88	Bel	Belgium	4	0.10
89	capt	captain	4	0.10
90	Dr.	Doctor	4	0.10
91	FAT	Football Association of Thailand	4	0.10
92	FC	Football Club	4	0.10
93	FIA	Federation of International Automobile	4	0.10
94	Ina	Indonesia	4	0.10
95	KOs	knockout	4	0.10
96	kph	kilometre per hour	4	0.10
97	Mar	Morocco	4	0.10
98	mins	minutes	4	0.10
99	Nor	Norway	4	0.10
100	NPC	National Provincial championship	4	0.10
101	NRL	National Rugby League	4	0.10
102	Por	Portugal	4	0.10
103	Svk	Slovakia	4	0.10
104	Ukr	Ukraine	4	0.10
105	WBA	World Boxing Association	4	0.10
106	ACB	Australian Cricket Board	3	0.07
107	Col	Colombia	3	0.07
108	Cro	Croatia	3	0.07
109	def	defeated	3	0.07
110	DQ	disqualify	3	0.07
111	FIBA	International Basketball Federation	3	0.07

Rank	Word	Full Name	Freq.	Percent
112	Hun	Hungary	3	0.07
113	IRB	International Rugby Board	3	0.07
114	ISIA	Ice Skating Institute of Asia	3	0.07
115	kms	kilometres	3	0.07
116	Lt	Lieutnant	3	0.07
117	LTAT	Lawn Tennis Association of Thailand	3	0.07
118	Maj	Major	3	0.07
119	Mas	Malaysia	3	0.07
120	Mex	Mexico	3	0.07
121	NFC	National Football Conference	3	0.07
122	Slo	Slovenia	3	0.07
123	TKO	technical knockout	3	0.07
124	UK	United Kingdom	3	0.07
125	WACA	Western Australian Cricket Association	3	0.07
126	WBC	World Boxing Council	3	0.07
127	WGC	World Golf Championship	3	0.07
128	wkt	wicket	3	0.07
129	WTC	World Trade Center	3	0.07
130	AMAT	Amateur Muay Thai Association of Thailand	2	0.05
131	Bah	Bahamas	2	0.05
132	Blr	Belarus	2	0.05
133	Bul	Bulgaria	2	0.05
134	CBA	China Basketball Association	2	0.05
135	HRH	His Royal Highness	2	0.05
136	hrs	hours	2	0.05
137	IFMA	International Federation of Muay Thai Amateur	2	0.05
138	IMG	International Management Group	2	0.05
139	Jam	Jamaica	2	0.05
140	Jr.	Junior	2	0.05
141	kg	kilogram	2	0.05
142	Lat	Latvia	2	0.05
143	MVP	Most Valuable Player	2	0.05
144	Oct	October	2	0.05
145	Pros	Professionals	2	0.05
146	QPR	Queens Park Ranger	2	0.05
147	Qtr	Qatar	2	0.05
148	TDA	Thailand Darts Association	2	0.05
149	TDs	touchdowns	2	0.05
150	TGA	Thai Golf Association	2	0.05
151	UAE	United Arab Emirates	2	0.05
152	Utd	United	2	0.05
153	Uzb	Uzbekistan	2	0.05
154	VAT	Volleyball Association of Thailand	2	0.05
155	WDF	World Darts Federation	2	0.05
156	Wis	Wales	2	0.05
157	AFC	Asian Football Confederation	1	0.02
158	AFL	Australian Football League	1	0.02
159	AGU	Australian Golf Union	1	0.02
160	AHL	American Hockey League	1	0.02
161	ARU	Australian Rugby Union	1	0.02
162	Aut	Austria	1	0.02
163	BCC	Bangkok Christian College	1	0.02
164	CFL	Canadian Football League	1	0.02
165	CGF	Commonwealth Games federation	1	0.02
166	Cub	Cuba	1	0.02
167	Cyp	Cyprus	1	0.02
168	Dec	December	1	0.02
169	Den	Denmark	1	0.02
170	DFB	German Football Association	1	0.02

Rank	Word	Full Name	Freq.	Percent
171	EBU	European Badminton Union	1	0.02
172	Est	Estonia	1	0.02
173	Eth	Ethiopia	1	0.02
174	FBI	Federal Bureau of Investigation	1	0.02
175	Fij	Fiji	1	0.02
176	ft	feet	1	0.02
177	GMT	Greenwich Mean Time	1	0.02
178	HK	Hong Kong	1	0.02
179	IAMTF	International Amateur Muay Thai Federation	1	0.02
180	ICC	International Cricket Council	1	0.02
181	IFF	International Fencing Federation	1	0.02
182	IOC	International Olympic Committee	1	0.02
183	Ire	Ireland	1	0.02
184	ISU	International Skating Union	1	0.02
185	KTB	Krung Thai Bank	1	0.02
186	MLB	Major League Baseball	1	0.02
187	Mrs.	Mistress	1	0.02
188	NC	North Carolina	1	0.02
189	NFA	Nigerian Football Association	1	0.02
190	Nzl	New Zealand	1	0.02
191	PABA	Pan Asia Boxing Association	1	0.02
192	Pacs	Pacifics	1	0.02
193	Swi	Switzerland	1	0.02
194	Twn	Taiwan	1	0.02
195	USOC	U.S. Olympic Committee	1	0.02
196	VIP	Very Important Person	1	0.02
197	VIPS	Very Important Persons	1	0.02
198	WBO	World Boxing Organization	1	0.02
Total			4,147	100

APPENDIX E

Sentence Structure Analysis

Simple Sentences

1. Fifa's main concern has always been money.
2. Fortunately, this spectacle was not repeated for the final.
3. Tournaments in Thailand depend on big sponsors without any consideration for the young players.
4. Terry will look after the 11 players on the field.
5. I could accurately predict the appearance of these kinds of fans.
6. Lleyton Hewitt (Australia) beat Mikhail Youzhny (Russia) 6-3 6-3 7-5.
7. At Cheltenham, Gloucestershire won by six wickets.
8. Beckham spotted Nico Vaesen off his line and chipped the City goalkeeper from 25 metres in the 73rd minute.
9. Rich Aurilia added a run-scoring single in the ninth and had two hits.
10. A disappointed Boonchu partly blamed his new running shoes and a muscle pull for his disappointing performance.
11. They also lost their Tri-Nations title to New Zealand.
12. Northern Hemisphere teams fill 11 of the 15 positions in rugby union's Zurich World XV for 2002.
13. Alan Henderson and Dion Glover both added 14 points to the Atlanta attack.
14. After the match, Torino's manager, Renzo Ulivieri, paid tribute to the Italian champions.
15. He picked the team up on his shoulders.

This material is reserved for educational use only, not allowed for commercial use.

Forbidden to modify the content, and cite the document when use.

16. Martin Lapointe's second period goal tied the game for Boston.
17. Stars and Stripes sailed itself into the doldrums at one point.
18. Anaheim's aggressiveness has pushed the Yankees to the brink of elimination in the first round.
19. Witschge curled the ball over the wall of Barcelona players and into the right corner of the goal.
20. Kenyan Michael Mukomo had to settle for third place in 2:38.
21. A total of 10 Asian players, including five Thais, took part in this year's tournament.
22. Tiger Woods was voted PGA Tour player of the year on Wednesday by the Golf Writers Association of America.
23. Marat Safin and Spanish favourites Juan Carlos Ferrero and Carlos Moya were all beginning their bids for crucial ranking points in second-round games later yesterday.
24. Raffarin has cleared some of his schedule for the denouncement tomorrow.
25. The 32-year-old forward originally suffered the injury in a 2-1 defeat to Valladolid on October 20.
26. Both teams finished the season at the top of their respective divisions with records of 13 wins and five losses.
27. The Springboks conceded a record 30-10 loss to France and their first defeat in 33 years to Scotland on the two previous weekends.
28. Alvin Williams, Morris Peterson and Voshon Lenard each scored 17 points for Toronto.
29. James Posey led the losing Nuggets with 16 points.
30. Warner suffered the injury in the first quarter of Sunday's 13-10 loss to the Dallas Cowboys.
31. Bret Boone and Mike Cameron hit solo homers in the fourth for the Mariners.
32. The Americans had also beaten their rivals in the first round.

33. On Friday, Collingwood opened the playoffs with an upset of first-place finishers Port Adelaide.
34. Robb Nen pitched a scoreless ninth for his 33rd save of the season.
35. Mauresmo expects a tough match against her 26 year-old opponent on Montreal's hard court.
36. In the featherweight category, Pranee Kaengmak defeated Australian Tiana Caverely.
37. In the light-flyweight class, Siriporn Sonplod outpointed Erin Linky of Canada.
38. Paola Suarez and Virginia Ruano Pascual were given the women's doubles world champions award.
39. Stephon Marbury had 18 points and eight assists.
40. We've won the three games.
41. Chris Osgood made 21 saves.
42. Whitney added two assists.
43. Ronaldo should not be named European Footballer of the Year.
44. Koy Detmer will replace McNabb.
45. He took some good shots.
46. The Eagles defence didn't allow a touchdown.
47. Johnson notches his 20th win.
48. It was his second 100-yard rushing game of the season.
49. In New York, Larry Brown of the Philadelphia 76ers will be the head coach of the 2004 US Olympic men's basketball team.
50. The loss of Warner is a tremendous blow for the defending NFC champions.

Compound Sentences

1. The six and two-mile starts are at the same place on the track and this probably caused the jockey's confusion.
2. I was just playing bad and I have no confidence.
3. Thousands of English fans have travelled to Bratislava for the game but no other incidents were reported.
4. Sarafan had leaned across Falbrav first but the winner had then barged into the US challenger just before the end.
5. I haven't played in a tournament for months and this will be good preparation for the next English season.
6. Leeds United arrives in Bangkok tomorrow morning and there will be a press conference at the Merchant Court Hotel at 2 pm on the same day.
7. Thailand's Thongchai Jaidee struggled in his first appearance in the British Open yesterday and at press time he was five over par after nine holes.
8. She played really well at the right points and I made a lot of mistakes.
9. Woods was last in the four-man field with US\$125,000, but he won four skins.
10. I hit a lot of good golf shots and I started out well.
11. Five Yankees have started Game One of a post-season series, and four of them have struggled so far.
12. Brett Favre and Donovan McNabb each played just three series, but neither was too impressive.
13. I didn't play for a few weeks and I'm really tired.
14. I didn't win the French Open, but I'm going to do my best at Wimbledon.

15. Sievers kicked first-half field goals of 37, 39, 44 and 25 yards, and McGahee's 1-yard TD run gave the Hurricanes a 19-3 halftime lead.
16. Jay Williams added 11 points for the Americans and Baron Davis had 10.
17. France may have launched their Euro 2000 qualifying campaign with a victory but their performance in Cyprus will hardly have silenced the critics.
18. Several athletes in the US and Russia have sports psychologists and that's part of the secret of their successes.
19. There was no wind on Thursday but poor old Tom could only manage a six-over 77.
20. Praprut levelled up again and the scores were tied at 3-3.
21. Pavol Demitra opened the scoring just 16 seconds into the game for St Louis and Al MacInnis had the other Blues goal.
22. The strong support from their home crowd was fantastic and their Dutch coach Guus Hiddink improved their team a lot.
23. They have been winning trophies in previous Skate Asia competitions and this year they are well prepared, hence they should do well.
24. The umpire reprimanded the boisterous crowd twice but excited screams for all four players peppered the match.
25. In closer matches yesterday, Geelong edged St. Kilda 16.6 (102) to 15.11 (101) and Melbourne defeated the West Coast Eagles 15.16 (106) to 15.10 (100).
26. A 15-year-old shooter earned Bangladesh its first medal and title and a 21-year-old weightlifter won three golds for Cameroon.
27. In more third-day action, top seed Albert Costa of Spain beat Costa Rican Juan Antonio Marin 6-1, 6-3, and Costa's countryman, No. 2 seed Juan Carlos Ferrero triumphed over Irakli Labadze of Georgia, 6-4, 6-4.

28. Walker later hit his homer off Tom Davey and Aaron Boone also had a home run for Cincinnati.
29. Resurgent England has five players in the side and Six Nations champions France provide three.
30. In Chicago, Magglio Ordonez hit a two-run homer in the sixth inning and the White Sox held on for a 3-2 triumph over the Boston Red Sox.
31. They both went for the tackle and he got the ball.
32. Injury-riddled Toronto dressed just nine players and Jelani McCoy did not see any action.
33. The Fighting Irish lost three of their seven fumbles and backup Pat Dillingham threw two interceptions.
34. On the second mark, it was just 16 seconds behind OneWorld but on the third mark Stars and Stripes was one minute 57 seconds behind.
35. The men's football competition will kick off on September 27th and the final will be on October 13th.
36. I spoke to Tyson's manager a few weeks ago about the possibility and he was quite amiable.
37. Iceland is at home and most of the players play in England with Premier League clubs or the First Division, or in Belgium and Sweden.
38. My swing is no good now and my game is no good.
39. Every win is worth one point and the top eight teams will reach the quarter-finals.
40. Osotspa has many rising stars and we cannot underestimate them at all.
41. The injuries are not considered life threatening and the prognosis for a long-term recovery is excellent.
42. We absolutely love big events and we will not fail you.
43. Actually, I wanted Choketawee Promrut but he is tied up with the S-league.

This material is reserved for educational use only, not allowed for commercial use.

Forbidden to modify the content, and cite the document when use.

44. I prefer four but we have had a lot of problems in the centre.
45. Thailand has often beaten Australia in the past few years, so Shanrit is not much worried about them.
46. Ferrari's domination has hit ticket sales for Sunday's US Grand Prix but the race will still draw the biggest one-day crowd of the Formula One season.
47. Michael Jordan's right knee made it through the exhibition season without any problems, and that's the most important thing for the Washington Wizards.
48. Scotland play hosts to Romania and Wales take on Fiji in Cardiff in other Tests.
49. The competition for the WDF AsiaPacific Cup starts today and continues until Thursday and on Friday and Saturday the Beer Chang Thailand Darts Open takes over.
50. Thailand kept up the pressure and Anucha made it 3-2 and then clinched his hat-trick in the final minute for a 4-2 scoreline.

Complex Sentences

1. Australia will be missing for the first time in 40 years and 14 Cup contests when America's Cup racing starts in Auckland next week. .
2. The number one pick in this year's NBA draft got a standing ovation when he was introduced before a pre-season game with the Orlando Magic.
3. Three-time Major winner Nick Price will grace the fairways of the US\$250,000 Macau Open this year when the tournament tees-off from October 17-20.
4. Ljubicic now moves on to a second round clash with Wimbledon semi-finalist Xavier Malisse after the Belgian battled past Julien Boutter 6-7, 7-6, 6-4.
5. San Antonio coach Gregg Popovich was ejected for picking up his second technical foul of the game as the first half ended.

6. Asloum scored a TKO over Saensak Singmanasak of Thailand two minutes into the fourth round of their super flyweight bout while Ouballi won by TKO over Nuafa Singjaiyai 50 seconds into the second round of their super featherweight fight.
7. In a way, Muay Thai is still behind traditional sports of certain Asian countries such as Japan's judo, Korea's taekwondo, China's wushu and India's kabaddi although it is far more exciting.
8. The winner of the Louis Vuitton Cup will take on holders Team New Zealand, who beat Prada 5-0 two years ago, in the 31st America's Cup match from February 15, 2003.
9. The season-ending Volvo Masters of Asia will be an event that holds great importance to the players.
10. Australian Bayliss has also joined the exodus to MotoGP, which did away with the old 500cc restriction to the top category of grand prix motorcycling before the start of last season.
11. The best player of each match at this year's Asian Rugby Championship, which starts today, will be presented with a special award.
12. Bobby Kielty also homered for the Twins, who won three of the four games in the series.
13. Amanda Coetzer is another experienced player who has not lost the hunger for competition.
14. OneWorld's semi-finals opponents are defending Louis Vuitton Cup challengers Prada of Italy, who are seeking OneWorld's disqualification for allegedly having boat design secrets taken from other teams.
15. Smyth, who became the oldest winner on the European Tour with victory in last year's Madeira Island Open at the age of 48, missed a birdie putt from seven feet at the last.
16. Former CART champion Juan Pablo Montoya says Brazilian Cristiano da Matta should take the plunge and follow him into Formula One.
17. South African promoter Rodney Berman said on Friday that he was in talks with Tyson's camp to bring the self-styled "baddest man on the planet" to the boxing-mad country.

This material is reserved for educational use only, not allowed for commercial use.

Forbidden to modify the content, and cite the document when use.

18. Higgins is still optimistic he can compete at next month's Embassy World Championship qualifiers at the Manhattan's Snooker Club in Harrogate.
19. While some athletes and officials are focusing on the Asiad, to be held in Busan, South Korea in September, others have been caught up in controversy.
20. My coach told me I am now a part of swimming trivia.
21. In July, Paradorn had told AFP that being the underdog made him more relaxed.
22. Please note that programme timing is liable to change without prior notice.
23. Thailand have been boosted by their performance in the recent World Grand Prix in which they impressively fought heavyweights like Cuba, Brazil and the US a tough time although the Thais lost all their games.
24. Ferguson, who would have been watching the coming soccer season from his armchair, hopes that the empty trophy cabinet will spur the team into even greater efforts.
25. Woosnam had previously said he wanted to lead Europe for the 2010 Ryder Cup match when it is played in his home country Wales at Celtic Manor.
26. Lewis, who turned 37 on Monday, confirmed his standing as the best heavyweight in the world on June 8 when he stopped Tyson in the eighth round.
27. Matt Williams tripled in three runs for Arizona, which is still tied with St Louis for wins as they battle for home field advantage in the first round of the playoffs.
28. Maleeva, who had beaten Williams in two of their four previous meetings, said she was confident going into the match.
29. Double major champion Sandy Lyle again showed he is gradually getting back the game that won him the 1985 British Open and 1988 US Masters, as he returned a 68 to be in a large group a shot off the pace.
30. Juventus manager Marcello Lippi has warned his club's rivals that the Italian champions are stronger than last year after they beat local rivals Torino 4-0 on Sunday to go top of Serie A.

31. World and Olympic champion Maurice Greene has vowed to regain the world 100 metres record next season.
32. Yao missed three weeks of training camp because of the Asian Games and has only 10 days to prepare for the NBA season, which the Rockets will start on October 30 against the Indiana Pacers.
33. The tournament will be played in four divisions on three rugby pitches with the Police Club/Stadium Complex, behind the Rama Garden Hotel on Vibhavadi Soi 60, providing the main venue.
34. After winning the Myanmar Open at the start of 2002, Thongchai has found that dividing his time between the highly-competitive Japanese tour and the Davidoff Tour has taken its toll.
35. Muay Thai would have better international recognition if there was only one ruling body.
36. The Roosters, led by former Test captain Brad Fittler, will find nothing new in playing to a sell-out crowd of 80,000 at Sydney's Olympic Stadium.
37. Based on leadup form, England have not shown enough to suggest that they will stretch Australia, who have not lost a home series at home since going down 2-1 to the West Indies nine years ago.
38. The quarter-final round matches are scheduled to be played today at the same venue, starting at 10 am.
39. No. 3 Ohio State (11-0) remained the only other major unbeaten team with a 10-6 win over Purdue, keeping the Buckeyes's hopes for a national title alive.
40. Competing in the heats of the 50 metres freestyle for disabled swimmers, she clearly had a severe lack of mobility in her legs.
41. By the end, the willowy Hantuchova, enjoying her breakthrough season on the WTA Tour, was clearly tired and making many unforced errors under the lights on the carpeted floor of the Schluefweg Arena.

42. Ranked No. 4 in the IBF world rankings, the 23-year-old Zhou picked up US\$7,500 and a gold medal for her victory.
43. Kostelic also won the combined event which twinned the slalom with Saturday's downhill, giving her an extra 100 points in the World Cup standings.
44. While Li easily outshone his nearest rivals to grab the top prize in both the vault and parallel bars, Russian Svetlana Khorkina failed in her bid to win the beam title with a scrappy display of her balancing skills.
45. Allen Iverson scored 29 points as the Philadelphia 76ers stayed unbeaten at home, downing the New Orleans Hornets 108-87 in NBA action on Monday.
46. Sergei Fedorov and Kirk Maltby both scored twice as the Detroit Red Wings made Curtis Joseph's debut in goal a successful one, beating the San Jose Sharks 6-3 on Thursday.
47. It is our pleasure for Thailand to host this international darts tournament for the first time.
48. This is a guide to sports showing on UBC and free TV this week.
49. Since Bowling Green has no realistic chance of climbing into the upper reaches of the BCS, there really are just three teams that have the inside track for those No. 1 and No. 2 spots.
50. There was no way back for an out-of-sorts Kuerten, with Paradorn completing victory in 76 minutes to set up a second-round clash with British No. 1 Tim Henman.

Compound-Complex Sentences

1. School knows that racing comes first with me and they've been really supportive.
2. I still have the feeling I'm improving so I like to play.
3. The manager has told us we are going for the title and I would find it hard to leave here without winning something.
4. I've played against Australia and I don't believe our failings are down to a mental thing.

5. Australia will be represented by top Suburban League Club Beecroft Cherrybrook which includes a number of ex-Eastwood players and the Club has every intention of being in the finish of the Champion-Ferodo Cup.
6. My game has been in good shape this year so it would mean a lot to me to take the Macau Open title at my first attempt.
7. We have 17 races on the calendar, and regrettably one of them will have to go to make way for Shanghai.
8. They have lost a couple of players and I think they have also had a few injuries but German sides are always very resilient.
9. I traveled a long way and I'm here to try and win the tournament.
10. He joined McRae in his junior days in the sport and the pair rose to the top, winning the 1995 world title on the final round in Britain.
11. The tournaments making up the circuit will be held at Krisda City in Nakhon Pathom from October 8-9, Palm Hill resort from November 23-24, Bonanza course in Nakhon Ratchasima from February 8-9, Lam Look Ka country course in Pathum Thani from March 27-28, Water Mill course in Nakhon Nayok from May 8-9 and the final tournament will be held at Panya Park in Suwintawong from June 19-20.
12. I have played golf in Asia before and I am really looking forward to a return I have not been before, Macau.
13. The Irish (8-1) gave the ball away all day, and the Eagles (5-3) gladly took advantage, scoring both touchdowns courtesy of first-half turnovers.
14. Fifth-seeded Paradorn had few alarms as he won his 14th match out of his last 16 on the ATP Tour and he now has a good chance of winning through to the final stages of the tournament.
15. Kazakhstan was awarded the silver medal when Ponomarev put up a good fight for his 117, but Sergey Yakshin (115) and Sergey Kolos (115) saw the total reach to 347.

16. When I lost the first set, I was really happy with the way I was fighting and I would not give up.
17. Spanish newspapers have been speculating that Brazilian World Cup winner Ronaldo could be the first name on Real's shopping list but Perez all ruled out a bid for the Inter Milan forward.
18. Even if you win a tournament, the next week you have to play again and you never get time to enjoy it.
19. Nagashima was a player who follows his own pace and I fought him by keeping a distance.
20. Thorpe, who holds world records over the 200-, 400- and 800-metre freestyle, gave Australia its sixth gold medal but the Americans drew level when world record-holder Natalie Coughlin eclipsed the race in the women's 100 backstroke in the second final of the night.
21. We've shown over the summer we can make some good totals but you can't hide away from the injuries.
22. I had the feeling I was a bit tight in my game and in my mind but it's a good experience.
23. It's very hot, but I'm happy to win.
24. Zito (22-5) fell behind 4-1 in the fourth inning, but Oakland scored six runs in the fourth to leave the two teams tied for first place with just 10 games left in the regular season.
25. Willis McGahee ran for 154 yards and a touchdown and Todd Sievers kicked four field goals to guide the University of Miami Hurricanes to a 26-3 defeat of Tennessee on Saturday.
26. Scott Spiezio drove in three runs and Ramon Ortiz pitched six solid innings to steer the Anaheim Angels to a 5-2 victory over the Oakland Athletics in the American League on Tuesday.
27. Alfonso Soriano stroked a pair of RBI singles, including his 200th hit of the season, and Jason Giambi blasted his 38th home run to help the New York Yankees to a 7-1 win over the Tampa Bay Devil Rays.

28. At Phoenix, Randy Johnson tossed seven shutout innings and Steve Finley hit a three-run homer to spark Arizona to a 12-0 romp over the Padres.
29. In Miami, Willis McGahee had 221 total yards, and the No. 1 Hurricanes extended their winning streak to 26 games by scoring three touchdowns during a three-minute span of the fourth quarter to beat Boston College 38-6.
30. We won't be able to play at a high level and we won't be able to perform well if there is no link between the defence and the attack.
31. Houston won for the first time this season and No. 1 draft pick Yao Ming got his first NBA points as the Rockets edged the Denver Nuggets 83-74 in NBA play on Friday.
32. Shaquille O'Neal had 35 points and 10 rebounds and Kobe Bryant just missed a triple-double as the LA Lakers won consecutive games for the first time in three weeks, downing the slumping Toronto Raptors 104-88 on Sunday.
33. Todd White had a goal and an assist and Marian Hossa tied for the league lead in goalscoring with his 19th of the season as the Ottawa Senators defeated the Boston Bruins 5-2 on Thursday.
34. In Denver, Chauncey Billups had 21 points and Richard Hamilton added 15 as the Detroit Pistons coasted to an 89-63 win over the struggling Nuggets.
35. In yesterday's match the players were tied 2-2 at the interval but Kobkit hit a purple patch as he took the next three frames.
36. In Oakland, California, Michael Olowokandi had 24 points and 10 rebounds and Elton Brand added 17 and 11 as the Los Angeles Clippers posted a 99-89 victory over the Golden State Warriors.
37. Midfield ace Juninho was having more and more influence on the game, and it was his cross from a free-kick which set up the winner for Mahamadou Diarra four minutes after the break.
38. Beckham crowned his first league start since November with a magnificent goal but it was Fabien Barthez who was credited with delivering the three points with two key second-half saves when United were only 1-0 up.

This material is reserved for educational use only, not allowed for commercial use.

Forbidden to modify the content, and cite the document when use.

39. Scolari was criticised by the Brazilian people before the tournament for not picking Romario in his squad but he proved that Brazil could win the World Cup without Romario.
40. It wasn't an easy race but I'm always confident that I can perform well.
41. Rusedski's serve was unstoppable throughout and he made the most of the lightning fast courts in Cincinnati, sending one delivery down at 135mph.
42. Jerry Stackhouse had 28 points for the Wizards and Jordan was still a factor, picking up nine assists and eight rebounds while playing a season high 40 minutes.
43. Quarterback Nate Hybl threw four interceptions, and Texas scored on them all, taking a 17-11 lead midway the third quarter Saturday.
44. In their first competitive match since a humiliating first-round exit at the World Cup, France had to come from behind to clinch a 2-1 win and coach Jacques Santini was highly critical of his team's defending.
45. Couples followed with a birdie to win the 11th and O'Meara made a birdie to capture the 12th.
46. In the other last 16 round matches yesterday, Noppadol Saengnil played brilliantly to defeat No. 5 ranked Somchai Kaosamran 4-1 and young Nithiwat Kanajanasri beat Pitak Santiniwate 4-2.
47. At the Australian Open this year she beat Williams to reach the semi-finals, and she was also a quarter-finalist at both the French Open and Wimbledon.
48. It took the 22-year-old American 18 minutes to get onto the scoreboard, but she held serve and took the Belgian's serve to love to draw level 2-2 following a double fault by Henin.
49. The athletes admitted that the windy conditions in Poland were good for their training but they complained about poor living conditions because they were short of funds as the Windsurfing Association of Thailand was too slow to transfer money to them from Thailand.
50. We never let our heads go down and the second half was much better.

BIOGRAPHY

- Name :** Mr. Mana Tosatjawong
- Date of Birth :** 20 March 1971
- Place of Birth :** Bangkok, Thailand
- Education :**
- 1990-1997 Bachelor's degree in Mass Communication, majoring in Mass Communication and minoring in Sociology and Anthropology
Ramkhamhaeng University
- Work Experience :**
- 1990-1999 Housekeeping Coordinator in Housekeeping Department at The Regent of Bangkok Hotel
- 1999-2003 Guest Service Agent in Room & Guest Service Department at Pathumwan Princess Hotel
- 2004-2005 Part-time teacher at SMIT International School
- 2006 Lecturer at Bangkok Suvarnabhumi College

