

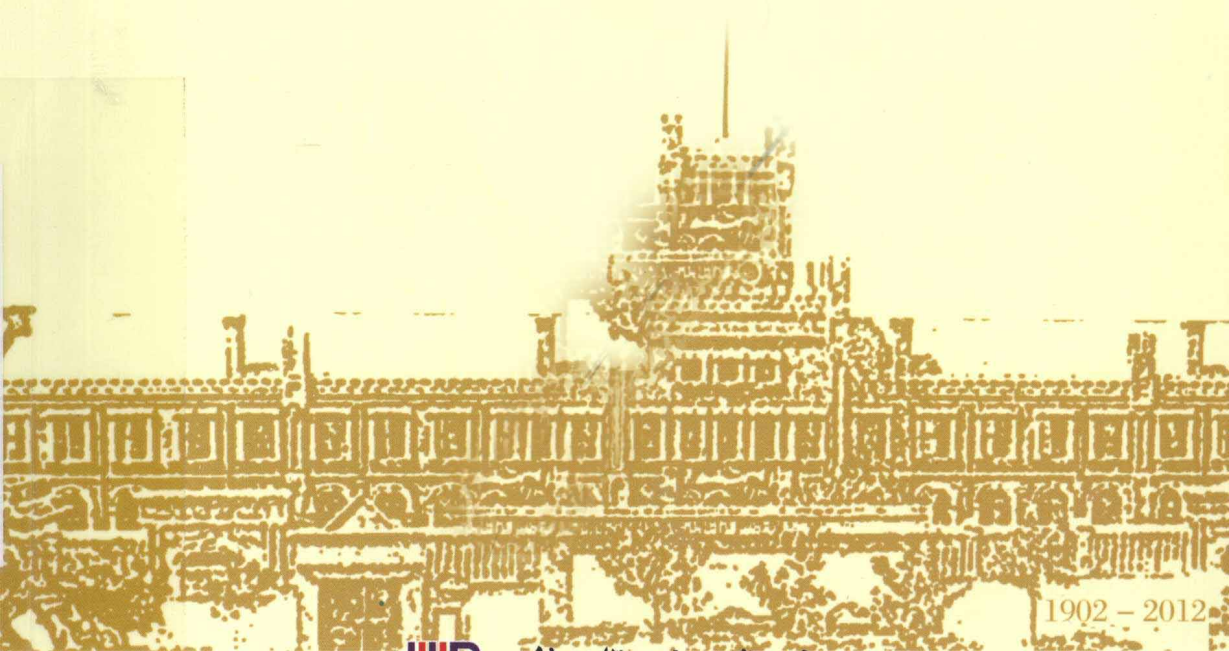


山西大学建校 110 周年学术文库

基于语料库的中国高级英语 学习者词块使用研究

A Corpus-based Analysis of the Use of Lexical Chunks
by Chinese Advanced EFL Learners

张 霞 著



1902 - 2012



科学出版社



教育部人文社会科学重点研究基地重大项目

基于语料库的中国高级英语 学习者词块使用研究

A Corpus-Based Analysis of the Use of Formulaic Chunks
by Chinese Advanced EFL Learners

李 西 西





西大学建校 110 周年学术文库

基于语料库的中国高级英语学习者词块使用研究

A Corpus-based Analysis of the Use of Lexical Chunks
by Chinese Advanced EFL Learners

张 霞 著

科学出版社

北 京

内 容 简 介

词块集中体现了语言使用中规约性与创新性的辩证统一,词块使用特征往往提示出文本构成的机制以及学习者的语言能力特征。本书设计了一套行之有效的词块抽取和赋码方案,并将之应用于中国高级英语学习者词块使用的调查分析,从结构功能和词汇语法两个层面考察了学习者和本族语者使用词块的趋势;还对学习者和本族语者、母语不同的学习者、学习者和教材用法进行了对比分析,旨在揭示学习者用法中产生偏差的内在机理,以期对教学有所启示。

本书适合高校英语教师、语言学类研究人员、高校在读语言学类研究生及对语言量化研究感兴趣的本科生阅读。

图书在版编目(CIP)数据

基于语料库的中国高级英语学习者词块使用研究/张霞著. —北京:科学出版社, 2012. 4

ISBN 978-7-03-033978-2

I. ①基… II. ①张… III. ①英语—学习方法—研究—中国
IV. ①H319.3

中国版本图书馆 CIP 数据核字(2012)第 060261 号

责任编辑:刘彦慧 王昌凤/责任校对:宣 慧

责任印制:赵德静/封面设计:李恒东 无极书装

联系电话:010-6401 9074 电子邮箱:liuyanhui@mail.sciencep.com

科 学 出 版 社 出 版

北京东黄城根北街 16 号

邮政编码:100717

<http://www.sciencep.com>

中国科学院印刷厂印刷

科学出版社编务公司排版制作

科学出版社发行 各地新华书店经销

*

2012 年 5 月第 一 版 开本: B5 (720×1000)

2012 年 5 月第一次印刷 印张: 13 3/4

字数: 345 000

定价: 48.00 元

(如有印装质量问题, 我社负责调换)

总 序

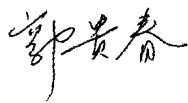
2012年5月8日，山西大学将迎来110年校庆。为了隆重纪念母校110年华诞，系统展现近年来山西大学创造的优秀学术成果，我们决定出版这套《山西大学建校110周年学术文库》。

山西大学诞生于“三千年未有之变局”的晚清时代，在“西学东渐，革故鼎新”中应运而生，开创了近代山西乃至中国高等教育的先河。百年沧桑，历史巨变，山西大学始终与时代同呼吸，与祖国共命运，进行了可歌可泣的学术实践，创造了令人瞩目的办学业绩。百年校庆以来，学校顺应高等教育发展潮流，以科学的发展理念引领改革创新，实现了新的跨越和腾飞，逐步成长为一所学科门类齐全、科研实力雄厚的具有地方示范作用的研究型大学，谱写了兴学育人的崭新篇章，赢得社会各界的广泛赞誉。

大学因学术而兴，因文化而繁荣。山西大学素有“中西会通”的文化传统，始终流淌着“求真至善”的学术血脉。不论是草创之初的中西两斋，还是新时期的多学科并行交融，无不展现着山大人特有的文化风格和学术气派。今天，我们出版这套丛书，正是传承山大百年文脉，弘扬不朽学术精神的身体力行之举。

《山西大学建校110周年学术文库》的编撰由科技处、社科处组织，将我校近10年来的优秀科研成果辑以成书，予以出版。我们相信，《山西大学建校110周年学术文库》对于继承与发扬山西大学学术精神，对于深化相关学科领域的研究，对于促进山西高校的学术繁荣，必将起到积极的推动作用。

谨以此丛书献给历经岁月沧桑、培育桃李芬芳的山大母校，祝愿母校在新的征程中继往开来，永续鸿猷。



二〇一一年十一月十日

Acknowledgements

This book is based on my doctoral dissertation. During the process of thesis writing and my study as a doctoral candidate in Shanghai Jiaotong University, I have been given various kinds of support and help.

First and foremost, I am deeply grateful to my supervisor Prof. Yang Huizhong for giving me the opportunity and helping me in many, many ways, even after his retirement, to develop my own ideas and the time to write them down. Heartfelt thanks must go to Prof. Wei Naixing for introducing me to the world of collocations and lexical chunks. A considerable part of the research presented here was inspired by or developed in projects led by him. Both of them took great pains in proofreading the drafts of the thesis and offered insightful guidance and valuable suggestions. Without support from them, this thesis would not see the light of day.

Special gratitude goes to Prof. Li Wenzhong, Prof. Pu Jianzhong and many, many others for their pioneering and inspiring work in the field of learner corpus research and studies of lexical chunks. My thanks are also extended to my various teachers and friends, linguistic and non-linguistic, for being my conscience and reminding me of the philosophy hidden in life from time to time.

Heartfelt thanks also go to the foreign researchers who unselfishly providing the precious literatures I needed, especially the Portuguese scholar Gael Dias for not only unselfishly providing me the source code of SENTA, but also translating the manual considerable in size from Portuguese to English when asked at short notice.

Special thanks go to my family for their support and patience on the many days when I was fighting writer's block, especially my husband for the support and help never missing and always there. Without them, none of this would have been possible. For me, the completion of the dissertation is both a journey of personality and a journey of knowledge, which shall be a treasure to draw upon for the rest of my life.

List of Abbreviations

ANOVA	Analysis of Variance
BNC	British National Corpus
CIA	Contrastive Interlanguage Acquisition
COLEC	Chinese Learner English Corpus
COLEN	The College English Textbook Corpus
COLSEC	College Learners' Spoken English Corpus
EFL	English as a Foreign Language
ELT	English Language Teaching
FLA	First Language Acquisition
FLOB	the Freiburg LOB Corpus of British English
LOCNESS	Louvain Corpus of Native English Essays
POS	Part of Speech
PV	Procedural Vocabulary
SLA	Second Language Acquisition
SOV	Subject-Object-Verb
SVO	Subject-Verb-Object
SVOC	Subject-Verb-Object-Complement
USAS	UCREL Semantic Analysis System
USE	the Uppsala Student English Corpus
VB	Visual Basic
XML	Extensible Markup Language

Contents

总序

Acknowledgements

Introduction1

Chapter 1 Lexical Chunks and Language Description6

- 1.1 The conventional use of language as a long-standing and long-recognized phenomenon6
 - 1.1.1 A de facto, long-standing and wide-spread phenomenon..... 6
 - 1.1.2 A long recognized phenomenon 9
- 1.2 The centralization of conventional language in neo-Firthian researches12
 - 1.2.1 The difference between the functionalist and mentalist paradigms..... 12
 - 1.2.2 From meaning and function to the conventional use of language..... 14
 - 1.2.3 How neo-Firthian arguments solve the problem of productivity..... 19
- 1.3 Lexical chunks as representing conventionality in language use21
 - 1.3.1 The nature of lexical chunks 21
 - 1.3.2 Features of lexical chunks..... 22
- 1.4 A review of the methodology development in studies of lexical chunks30
 - 1.4.1 The marginalization of phraseology and the salience approach..... 31
 - 1.4.2 The centralization of phraseology and the statistical approach..... 32
 - 1.4.3 Altenberg, Biber and the study of lexical chunks..... 35
- 1.5 Summary37

Chapter 2 Lexical Chunks and Language Acquisition39

- 2.1 A comparison.....39
- 2.2 Holistic language processing and FLA40
 - 2.2.1 Explanatory adequacy and the alternative expediency..... 41
 - 2.2.2 Holistic language processing and the route of FLA..... 42
 - 2.2.3 The role of holistic language processing in FLA 44
- 2.3 Lexical chunks and SLA45
 - 2.3.1 Differences and similarities between advanced EFL learners and first language acquiring children..... 46

2.3.2	Lexical chunks and the route of SLA	47
2.3.3	Lexical chunks and the success of SLA—the social-functional level	48
2.3.4	Lexical chunks and the success of SLA—the lexico-grammatical level.....	50
2.3.5	Lexical chunks and the success of SLA—extant studies.....	51
2.4	Procedural vocabulary and lexical chunks	53
2.4.1	The conception of procedural vocabulary	53
2.4.2	The relationship between procedural vocabulary and lexical chunk	55
2.5	Deviations in learners' use of lexical chunks—description and diagnosis	57
2.5.1	Two kinds of deviations and two approaches to the diagnosis	58
2.5.2	The CIA approach and the diagnosis of learner deviations.....	59
2.5.3	The research questions of the present research	61
2.6	The acquisition of lexical chunks in the classroom context.....	63
2.7	Summary	66
Chapter 3	The Methodology	68
3.1	An NLP perspective of the study of lexical chunks	68
3.1.1	Hybrid algorithms	69
3.1.2	Association measures.....	71
3.1.3	Statistical algorithms to decide MWU boundary	73
3.1.4	Dispersion.....	75
3.1.5	The extraction of procedural vocabulary	76
3.1.6	The automatic extraction of patterns around a certain word	78
3.2	The methodology of the present study	79
3.2.1	Corpora used—size, sampling and preprocessing.....	79
3.2.2	The extraction of lexical chunks: length range, frequency, dispersion and boundary determination.....	81
3.2.3	The filtering process.....	84
3.2.4	Tools of lexical chunk extraction	85
3.2.5	Annotation.....	86
3.2.6	The extraction of PV	88
3.2.7	Word-form based lexical chunks and POS-tag sequences around PV	90
3.3	The research design of the present study	93
Chapter 4	Word-form Based Lexical Chunks.....	95
4.1	The proportion issue.....	95

4.2	Types and tokens of lexical chunks in different length categories.....	96
4.3	Types, tokens and type/token ratio of lexical chunks in different corpora.....	98
4.4	The structure and function characteristics of lexical chunks	101
4.4.1	Dependent clause	107
4.4.2	Independent clause	110
4.4.3	Single clause constituent.....	111
4.4.4	Incomplete phrase	112
4.4.5	Multiple clause constituents.....	113
4.4.6	The interaction between function and corpus	123
4.5	Discussion and summary.....	126
	Appendix: Complementary statistical analyses.....	127
Chapter 5	POS-tag Sequences Around PV	131
5.1	Procedural words in the four corpora.....	131
5.1.1	Rank correlation analysis	131
5.1.2	Distribution of content and function words	135
5.1.3	Distribution of PV of various parts of speech.....	137
5.2	The relationship between procedural words and lexical chunks	140
5.2.1	The proof of the close relationship between PV and lexical chunks.....	140
5.2.2	Descriptive statistics.....	144
5.3	procedural words common to the corpora under study	146
5.3.1	Analysis of <i>take</i>	147
5.3.2	Analysis of <i>make</i>	156
5.3.3	Analysis of <i>time</i> and <i>way</i>	163
5.4	Summary	168
	Appendix: Complementary statistical analysis	169
Chapter 6	Major Findings Revisited.....	173
6.1	Major findings and conclusions	173
6.2	Limitations and prospects.....	175
6.2.1	Higher degree of abstractness.....	175
6.2.2	Semantic prosody and discoursal functions	176
6.2.3	The developmental dimension	177
6.2.4	Procedural vocabulary.....	178
	References	180
	Appendix	191

List of Figures

Figure 1-1	Barlow's (2005) diagram of schema-based grammar.....	20
Figure 1-2	Van Lancker's continuum figure (cited from Wray, 2002: 64).....	30
Figure 1-3	an illustration of the methodology development of phraseology study.....	30
Figure 1-4	the linear distribution of multiple clause constituents (Altenberg, 1998: 111).....	36
Figure 2-1	relative proportions of holistic and analytical involvement in language processing from birth to adulthood (schematic representation) (Wray and Perkins, 2000).....	43
Figure 2-2	Contrastive Interlanguage Analysis (Granger, 1996).....	59
Figure 3-1	lexical chunk structure diagram in the present study	86
Figure 3-2	lexical chunk function diagram in the present study.....	87
Figure 3-3	the multiple CIA comparisons conducted.....	94
Figure 4-1	proportions of lexical chunks in four corpora	95
Figure 4-2	lexical chunk types in four corpora.....	97
Figure 4-3	lexical chunk tokens in four corpora.....	97
Figure 4-4	lexical chunk types collapsed by structure in two spoken corpora	98
Figure 4-5	multiple clause constituents collapsed by distribution in two spoken corpora.....	98
Figure 4-6	stems collapsed by function in two spoken corpora	98
Figure 4-7	lexical chunk types of the four corpora	99
Figure 4-8	lexical chunk type clustering of the four corpora.....	99
Figure 4-9	lexical chunk tokens of the four corpora	99
Figure 4-10	lexical chunk token clustering of the four corpora.....	99
Figure 4-11	TTR of the four corpora	100
Figure 4-12	inversed TTR of the four corpora	100
Figure 4-13	TTR clustering of the four corpora	100
Figure 4-14	TTR clustering of the four corpora after excluding 2-word chunks	100
Figure 4-15	inversed TTR clustering of the four corpora after excluding 2-word chunks.....	101
Figure 4-16	TTR of chunks of different lengths in the four corpora.....	101
Figure 4-17	lexical chunk types collapsed by structure in two spoken corpora	102
Figure 4-18	MCC types collapsed by distribution in two spoken corpora	102
Figure 4-19	MCC tokens collapsed by distribution in two spoken corpora	102

Figure 4-20	stem tokens collapsed by function in two spoken corpora	103
Figure 4-21	lexical chunks types collapsed by structure in two written corpora.....	104
Figure 4-22	lexical chunk tokens collapsed by structure in two written corpora	104
Figure 4-23	MCC tokens collapsed by distribution in two written corpora.....	104
Figure 4-24	stem tokens collapsed by function in two written corpora	105
Figure 4-25	medial tokens collapsed by function in two written corpora	105
Figure 4-26	lexical chunk types collapsed by structure in four corpora (proportion).....	106
Figure 4-27	lexical chunk tokens collapsed by structure in four corpora (proportion).....	106
Figure 4-28	MCC types collapsed by distribution in four corpora (proportion)	106
Figure 4-29	MCC tokens collapsed by distribution in four corpora (proportion)	106
Figure 4-30	dependent clause types in four corpora	107
Figure 4-31	dependent clause tokens in four corpora	107
Figure 4-32	clustering of the four corpora by dependent clause tokens	107
Figure 4-33	dependent clause types and tokens collapsed in four corpora (proportion)	108
Figure 4-34	dependent clause types collapsed in four corpora	108
Figure 4-35	dependent clause tokens collapsed in four corpora.....	108
Figure 4-36	clustering of the four corpora by dependent clause type distribution	108
Figure 4-37	clustering of the four corpora by dependent clause token distribution	109
Figure 4-38	independent clause types and token collapsed in two spoken corpora (proportion).....	110
Figure 4-39	independent clause types collapsed in two spoken corpora.....	110
Figure 4-40	independent clause tokens collapsed in two spoken corpora	110
Figure 4-41	the function diagram for MCCs, SCCs and IPs	111
Figure 4-42	SCC types and tokens collapsed by function in four corpora (proportion)	112
Figure 4-43	referential SCC types collapsed in four corpora (proportion)	112
Figure 4-44	attributive SCC types collapsed in four corpora (proportion).....	112
Figure 4-45	incomplete phrase types collapsed by function in four corpora	113
Figure 4-46	frame types including wobbly springboards in four corpora.....	114
Figure 4-47	frame tokens including wobbly springboards in four corpora	114
Figure 4-48	frame types excluding wobbly springboards in four corpora	114
Figure 4-49	frame tokens excluding wobbly springboards in four corpora	114
Figure 4-50	frame types and tokens collapsed by function in four corpora (proportion).....	115
Figure 4-51	onset types in four corpora.....	115
Figure 4-52	onset tokens in four corpora.....	115
Figure 4-53	onset types and tokens collapsed by function in four corpora (proportion).....	116

Figure 4-54	stem types in four corpora.....	116
Figure 4-55	stem tokens in four corpora.....	116
Figure 4-56	stem types and tokens collapsed by function in four corpora (proportion).....	117
Figure 4-57	medial types in four corpora.....	117
Figure 4-58	medial tokens in four corpora.....	117
Figure 4-59	medial types and tokens collapsed by function in four corpora (proportion)	117
Figure 4-60	rheme types in four corpora.....	118
Figure 4-61	rheme tokens in four corpora.....	118
Figure 4-62	rheme types and tokens collapsed by function in four corpora (proportion)	119
Figure 4-63	organizational types collapsed by distribution in four corpora	120
Figure 4-64	organizational tokens collapsed by distribution in four corpora	120
Figure 4-65	MCC function types collapsed by distribution in the learner spoken corpus	120
Figure 4-66	MCC function tokens collapsed by distribution in the learner spoken corpus	121
Figure 4-67	MCC function types collapsed by distribution in the native spoken corpus	121
Figure 4-68	MCC function tokens collapsed by distribution in the native spoken corpus	121
Figure 4-69	MCC function types collapsed by distribution in the learner written corpus	121
Figure 4-70	MCC function tokens collapsed by distribution in the learner written corpus	121
Figure 4-71	MCC function types collapsed by distribution in the native written corpus	121
Figure 4-72	MCC function tokens collapsed by distribution in the native written corpus	122
Figure 4-73	referential types collapsed by distribution in four corpora	122
Figure 4-74	referential tokens collapsed by distribution in four corpora	122
Figure 4-75	stance types collapsed by distribution in four corpora.....	122
Figure 4-76	stance tokens collapsed by distribution in four corpora	122
Figure 4-77	interactional types collapsed by distribution in four corpora	123
Figure 4-78	interactional tokens collapsed by distribution in four corpora	123

Figure 4-79	lexical chunk types collapsed by function in four corpora (left)	124
Figure 4-80	lexical chunk tokens collapsed by function in four corpora (right)	124
Figure 4-81	correspondence analysis of lexical chunk functions and corpora	125
Figure 4-82	two-dimension display of the distances between four corpora.....	125
Figure 4a-1	SCC types collapsed in four corpora.....	127
Figure 4a-2	SCC tokens collapsed in four corpora.....	127
Figure 4a-3	clustering of the four corpora by SCC type distribution	127
Figure 4a-4	clustering of the four corpora by SCC token distribution	128
Figure 4a-5	SCC referential types collapsed in four corpora.....	128
Figure 4a-6	SCC referential tokens collapsed in four corpora.....	128
Figure 4a-7	clustering of the four corpora by SCC referential type distribution	128
Figure 4a-8	clustering of the four corpora by SCC referential token distribution	128
Figure 4a-9	SCC attributive types collapsed in four corpora	129
Figure 4a-10	SCC attributive tokens collapsed in four corpora	129
Figure 4a-11	clustering of the four corpora by SCC attributive type distribution.....	129
Figure 4a-12	clustering of the four corpora by SCC attributive token distribution.....	129
Figure 4a-13	IP types collapsed in four corpora.....	129
Figure 4a-14	IP tokens collapsed in four corpora.....	130
Figure 4a-15	clustering of the four corpora by IP type distribution	130
Figure 4a-16	clustering of the four corpora by IP token distribution.....	130
Figure 5-1	clustering of the four corpora by top 100 PV rank	132
Figure 5-2	clustering of the six corpora by top 100 PV rank.....	134
Figure 5-3	PV collapsed by traditional parts of speech.....	135
Figure 5-4	proportions of PV in the database of word form based lexical chunks and the overall text of the four corpora	141
Figure 5-5	the averages of PV's lexical chunk use proportions in four corpora	142
Figure 5-6	the multiple CIA comparisons conducted.....	147
Figure 5a-1	clustering of the four corpora by rank correlation distance anchoring to BNC PV rank.....	170
Figure 5a-2	clustering of the four corpora by rank correlation distance anchoring to COLEC PV rank.....	170
Figure 5a-3	clustering of the four corpora by rank correlation distance anchoring to LOCNESS PV rank.....	171
Figure 6-1	formulaic sequences at different stages of acquisition and their future destination (Wray, 2002: 123)	178

List of Tables

Table 1-1	pattern flow, string, and loop (Hunston and Francis, 2000: 216)	20
Table 1-2	some representative three-way taxonomies of multi-word units and their correspondence with Halliday's three meta-functions	29
Table 1-3	an illustration of Altenberg's (1998) structure taxonomy and examples	35
Table 3-1	4-word lexical chunks in COLSEC filtered out.....	84
Table 3-2	an illustration of the database of word-form based lexical chunks with annotations (part of BNC database)	88
Table 3-3	an illustration of the database of PV with statistical information (part of LOCNESS database).....	89
Table 3-4	an illustration of the entries in the database of word-form based lexical chunks around PV	91
Table 3-5	COLSEC tag sequences of <i>take</i>	92
Table 4-1	6-word lexical chunks in four corpora.....	96
Table 4-2	top 10 stem types in two spoken corpora.....	103
Table 4-3	top 10 stem types in two written corpora.....	105
Table 4-4	top 10 infinitive types in four corpora (F: frequency).....	109
Table 4-5	rheme chunk examples from COLSEC.....	119
Table 4-6	statistical information of the correspondence analysis.....	125
Table 5-1	rank correlation analysis of the four corpora by top 100 PV rank	132
Table 5-2	rank correlation analysis of the six corpora by top 100 PV rank.....	133
Table 5-3	statistical information of content and function PV in four corpora	135
Table 5-4	the 2*2 ANOVA factoring in user (learner and native) and word type (content and function words).....	136
Table 5-5	the 4*2 ANOVA factoring in corpus and word type	137
Table 5-6	multiple comparisons between four corpora as to the distribution pattern of content and function words in PV	137
Table 5-7	PV common to at least three corpora.....	138
Table 5-8	PV unique to specific types of use or user	139
Table 5-9	proportions of PV in the word form-based chunk database and the overall text of the four corpora.....	141

Table 5-10	contingency table for COLSEC procedural words.....	141
Table 5-11	chi-square tests of PV's distribution in the word-form based chunk database and overall text of COLSEC.....	142
Table 5-12	one-way ANOVA of corpus effect upon word-form based chunk use of PV	143
Table 5-13	multiple comparisons between the four corpora as to the word-form based chunk use proportion of PV	143
Table 5-14	two-way ANOVA factoring in corpus and word type (content or function words)	144
Table 5-15	two-way ANOVA factoring in corpus and part of speech	145
Table 5-16	multiple comparisons among word-form based lexical chunk uses of PV of different parts of speech	145
Table 5-17	word-form based chunk use of <i>take</i> in four corpora	147
Table 5-18	typical word-form based chunks of <i>take</i> in four corpora.....	148
Table 5-19	qualified POS tag sequences around <i>take</i> and their typical examples in COLSEC	149
Table 5-20	qualified POS tag sequences around <i>take</i> and their typical examples in BNC	150
Table 5-21	qualified POS tag sequences around <i>take</i> and their typical examples in COLEC	151
Table 5-22	qualified POS tag sequences around <i>take</i> and their typical examples in LOCNESS.....	151
Table 5-23	qualified POS tag collocates around <i>take</i> in four corpora	152
Table 5-24	significant examples of tag collocates in COLSEC.....	152
Table 5-25	significant examples of tag collocates in BNC.....	152
Table 5-26	USE tag sequences of <i>take</i>	153
Table 5-27	USE tag collocates of <i>take</i>	153
Table 5-28	COLEN tag sequences of <i>take</i>	154
Table 5-29	COLEN tag collocates of <i>take</i>	154
Table 5-30	FLOB tag sequences of <i>take</i>	155
Table 5-31	FLOB tag collocates of <i>take</i>	155
Table 5-32	statistics of <i>take care of</i> and <i>take part in</i> in six corpora.....	155
Table 5-33	word-form based chunk use of <i>make</i> in four corpora	156
Table 5-34	typical word-form based chunks of <i>make</i> in four corpora.....	157
Table 5-35	significant tag-sequences around <i>make</i> in COLSEC.....	157

Table 5-36	significant tag collocates around <i>make</i> in COLSEC	158
Table 5-37	significant tag-sequences around <i>make</i> in BNC	158
Table 5-38	significant tag collocates around <i>make</i> in BNC	158
Table 5-39	significant tag-sequences around <i>make</i> in COLEC	159
Table 5-40	significant tag collocates around <i>make</i> in COLEC	159
Table 5-41	significant tag-sequences around <i>make</i> in LOCNESS	160
Table 5-42	significant tag collocates around <i>make</i> in LOCNESS	160
Table 5-43	significant tag sequences around <i>make</i> in FLOB	161
Table 5-44	significant tag collocates around <i>make</i> in FLOB	161
Table 5-45	significant tag sequences around <i>make</i> in COLEN	162
Table 5-46	significant tag collocates around <i>make</i> in COLEN	162
Table 5-47	significant tag sequences around <i>make</i> in USE	162
Table 5-48	significant tag collocates around <i>make</i> in USE	163
Table 5-49	word-form based chunk use of <i>time</i> in four corpora	163
Table 5-50	typical word-form based chunks of <i>time</i> in four corpora	164
Table 5-51	context of <i>time to</i>	164
Table 5-52	context of <i>the time</i>	164
Table 5-53	significant tag sequences of <i>time</i> in COLSEC	164
Table 5-54	significant tag sequences of <i>time</i> in BNC	165
Table 5-55	significant tag collocates of <i>time</i> in COLSEC	165
Table 5-56	significant tag collocates of <i>time</i> in BNC	165
Table 5-57	significant tag sequences of <i>time</i> in FLOB	165
Table 5-58	word-form based lexical chunks of <i>way</i> in four corpora	166
Table 5-59	typical word-form based chunks of <i>way</i> in the corpora	166
Table 5-60	significant tag sequences of <i>way</i> in COLSEC	167
Table 5-61	significant tag collocates of <i>way</i> in COLSEC	167
Table 5-62	significant tag sequences of <i>way</i> in BNC	167
Table 5-63	significant tag collocates of <i>way</i> in BNC	168
Table 5a-1	rank correlation of the four corpora referring to BNC PV rank	169
Table 5a-2	rank correlation of the four corpora referring to COLEC PV rank	170
Table 5a-3	rank correlation of the four corpora referring to LOCNESS PV rank	171
Table 5a-4	contingency table for BNC procedural words	171
Table 5a-5	chi-Square test of BNC procedural words	171
Table 5a-6	contingency table for COLEC procedural words	172
Table 5a-7	chi-Square test of COLEC procedural words	172