

DOCTORAL THESIS

A corpus-based study of the forms and functions of BE in the interlanguage grammars of Chinese learners of English

Zhang, Yanyan

Date of Award:
2008

[Link to publication](#)

General rights

Copyright and intellectual property rights for the publications made accessible in HKBU Scholars are retained by the authors and/or other copyright owners. In addition to the restrictions prescribed by the Copyright Ordinance of Hong Kong, all users and readers must also observe the following terms of use:

- Users may download and print one copy of any publication from HKBU Scholars for the purpose of private study or research
- Users cannot further distribute the material or use it for any profit-making activity or commercial gain
- To share publications in HKBU Scholars with others, users are welcome to freely distribute the permanent URL assigned to the publication

**A Corpus-based Study of the Forms and Functions of BE in the
Interlanguage Grammars of Chinese Learners of English**

ZHANG Yanyan

**A thesis submitted in partial fulfilment of the requirements
for the degree of
Doctor of Philosophy**

Principal Supervisor: Prof. Tony HUNG

Hong Kong Baptist University

January 2008

Abstract

The English verb BE, with its complicated forms and multiple functions, has posed great learning difficulties for learners of English both as the first and a second language. Most previous studies on the acquisition of BE have only examined a small number of participants and used data from experimental tasks, which have failed to provide a full picture of how BE is acquired and used by learners. By adopting a corpus-based approach, this study presents a fuller picture of the forms and functions of BE in the interlanguage grammars of Chinese learners of English.

Two learner corpora were used in this research, namely, Hong Kong Baptist University Interlanguage Corpus (HKBUILC) and Chinese Learner English Corpus (CLEC), consisting of English essays written by students from Hong Kong and Mainland China, respectively. Three subcorpora of CLEC (i.e. ST2, ST3, ST4) were adopted, representing Chinese learners at three different English proficiency levels. To obtain a comprehensive distribution of interlanguage BE, all the tokens of BE, including non-target-like, overgenerated BE, as in “*Mary was lost the necklace she borrowed from her friend*”, and omission of BE, as in “*My mother would \emptyset angry with me*”, were annotated with respect to their syntactic and semantic properties. The data were then retrieved type by type for examination by using WordSmith, a concordancing software.

To account for the variable use of BE in learners’ interlanguage grammars, this study further investigated a few factors that may influence the production and omission of BE, such as the following linguistic distinctions: (1) copular vs. auxiliary BE; (2) finite (*am, is, are, was, were*) vs. nonfinite (*be, being, been*) BE; (3) stage-level vs. individual-level predicate types (Carlson, 1977); (4) different post-BE constituents, e.g. ergative vs. unergative verbs (Perlmutter, 1978); and (5) the differences between L1 and L2 (i.e. English BE vs. Mandarin *Shi*/Cantonese *Haih*).

The concordancing analysis leads to a comprehensive picture of the distributional patterns of BE in Chinese-English interlanguage grammars, with detailed information on the frequencies and rates of both non-target-like and target-like uses of every form and function of BE in context. Based on this, the in-depth cross-sectional (ST2 vs. ST3 vs. ST4) and cross-varietal (ST3 vs. HKBUILC) investigations into the variable use of BE, particularly omission and overgeneration, demonstrate that there exist grammatical contingences involving all the aforementioned factors. The findings from the learner corpora are further supported by additional data from experimental tasks. Two interlanguage rules are proposed to account for omission of BE and overgeneration of BE respectively.

(1) BE \rightarrow \emptyset / [+Asp]

(2) \emptyset \rightarrow BE/ _VP

and the English Department, especially Julie, Heitao, and Suan, who have kept me company on many quiet weekends and holidays in our offices with a long corridor. I would also like to thank Patricia, who has helped me in the collection of experimental data, and Uli, my officemate, who has brought me numerous relaxing moments in our communication auf Deutsch.

Last but not the least, my forever thanks and regrets go to my family, to whom I have owed so much. In the four years of separation, my husband, Jim, has showed his strongest support for my study with endless love and incredible patience. My parents and parents-in-law have given me great support and understanding for my pursuit of knowledge in linguistics.

I hope, with all my heart and soul, that I will not disappoint all those who have shown concern and care for me with this piece of work, and never will in years to come.

Table of Contents

DECLARATION	i
Abstract	ii
Acknowledgements	iv
Tables	x
Charts	xi
Tags Used for Data Coding.....	xii
Chapter 1. Introduction	1
1.1 Motivation for this Study	2
1.2 Purpose and Significance of the Study.....	5
1.3 Organization of the Dissertation.....	6
Chapter 2. Research Background	8
2.1 The Development of the Learner Language Theories	8
2.1.1 The Contrastive Analysis Hypothesis.....	9
2.1.2 The Error Analysis Hypothesis.....	10
2.1.3 The Interlanguage Hypothesis	11
2.1.4 Summary	13
2.2 Methodological Background	13
2.2.1 Theory-driven Approaches to Interlanguage Analysis.....	15
2.2.2 Data-driven Approaches to Interlanguage Analysis	18
2.3 An Integrative Approach Adopted.....	21
Chapter 3. Contrastive Analysis of English BE and Chinese Shi/Haih	25
3.1 The Forms and Functions of BE in Modern English Grammar	25
3.1.1 BE in Descriptive English Grammar	25
3.1.2 Linguistic Analyses of BE.....	26
3.2 The Forms and Functions of <i>Shi/ Haih</i> in Modern Chinese	30

3.3 A Comparison between English BE and Chinese Shi/Haih.....	34
Chapter 4. Previous Studies on the Acquisition of BE	37
4.1 Morpheme Studies on BE in Language Acquisition	37
4.1.1 Morpheme Studies on BE in First Language Acquisition	37
4.1.2 Morpheme Studies on BE in Second Language Acquisition.....	39
4.1.3 Summary of the Morpheme Studies on BE	40
4.2 To Be or Not to BE: the Variable Use of BE	40
4.2.1 The Variable Use of BE in Child L1 English	41
4.2.2 The Variable Use of BE in World Englishes.....	45
4.2.3 The Variable Use of BE in L2 English	49
4.3 Previous Studies on BE in Chinese-English Interlanguages	58
4.3.1 Omission of BE in Chinese-English Interlanguages	58
4.3.2 Overgeneration of BE in Chinese-English Interlanguages.....	59
4.3.3 Systematic Variations of BE in Chinese-English Interlanguages	60
4.4 Summary of Previous Research on BE in Language Acquisition	61
Chapter 5. Research Method	65
5.1 Research Questions	65
5.2 Databases.....	67
5.2.1 Learner Corpora of Chinese Learners of English.....	67
5.2.2 Standard English Corpus	69
5.2.3 Experimental Data.....	71
5.3 Computer-aided Data Processing	73
5.3.1 Data Coding Task	73
5.3.2 Data Coding Procedures.....	87
5.4 Data Analysis Procedures.....	91
Chapter 6. Results of the Present Study	96
6.1 Distributional Patterns of Overt BE in Chinese-English Learner Corpora	96
6.1.1 Detailed Distributions of Overt BE in the Learner Corpora.....	99
6.1.2 Comparison of Distributions of BE in Native and the Learner Corpora	102

6.1.3 Summary of the Distribution of BE.....	106
6.2 Distributional Patterns of Null BE in Chinese-English Learner Corpora	107
6.2.1 Overall Distribution of Null BE in the Learner Corpora.....	108
6.2.2 Distribution of Null Auxiliary BE in contrast to Pseudo-Passive Cases	109
6.2.3 Summary of the Distribution of Null BE.....	110
6.3 The Variable Use of BE in Chinese-English Learner Corpora.....	111
6.3.1 Major Types of Misuses of BE in the Learner Corpora.....	111
6.3.2 Other Variable Uses of BE in Chinese-English Learner Corpora.....	113
6.4 Omission of Copular BE in Chinese-English Learner Corpora	117
6.4.1 Omission of Copular BE and Predicate Semantics.....	117
6.4.2 Omission of Copular BE and Syntactic Environments.....	120
6.4.3 Summary of Copular Omission	125
6.5 Overgeneration of BE in Chinese-English Learner Corpora.....	126
6.5.1 Token Distribution of Overgenerated BE	126
6.5.2 Rate Distribution of Overgenerated BE.....	127
6.5.3 Overgeneration of Finite BE vs. Nonfinite BE.....	129
6.5.4 Summary of Overgeneration of BE	129
6.6 Results from Experimental Tasks.....	131
6.6.1 Translation Task.....	131
6.6.2 Grammatical Judgement Task.....	133
6.6.3 Preference Test	136
6.7 Summary of the Results	138
Chapter 7. Discussion	141
7.1 Explaining the Distribution of BE.....	141
7.1.1 Higher Ratio of BE.....	141
7.1.2 More Frequent Use of Copular BE.....	143
7.1.3 More Frequent Use of Finite BE	144
7.1.4 Ranking of Copular Predicates	146
7.1.5 Summary on the Distribution of BE.....	146
7.2 Accounting for Omission of BE.....	147
7.2.1 Constraints on Omission of BE	149
7.2.2 Problems of the Previous Accounts.....	153

7.2.3 A Unified Account Proposed for Omission of BE	158
7.2.4 Accounting for the Difference between L1 and L2 Acquisition of BE	175
7.2.5 Summary on Omission of BE	178
7.3 Accounting for Overgeneration of BE	179
7.3.1 Overgeneration of BE and Overpassivization	181
7.3.2 Functions of Overgenerated BE	194
7.3.3 Interlanguage Rule on Overgeneration of BE	201
7.3.4 Summary on Overgeneration of BE	203
7.4 A Unified Account for Omission and Overgeneration of BE	204
7.5 Summary of Chapter 7	209
Chapter 8. Conclusion	211
8.1 Findings and Conclusions	211
8.2 Implications of this Study	215
8.3 Limitations of this Study	220
Appendices	223
Appendix 1. Translation Task	223
Appendix 2. Grammaticality Judgement Task	224
Appendix 3. Preference Test	226
Appendix 4. Distribution of BE with Contractions	227
Appendix 5. Detailed Results from Grammaticality Judgement Task	229
Appendix 6. Detailed Results from Preference Test	230
References	231
CURRICULUM VITAE	243