

**A CORPUS-BASED CONTRASTIVE ANALYSIS OF
SADNESS EXPRESSIONS IN ENGLISH AND CHINESE**

ZHANG RUIHUA

NATIONAL UNIVERSITY OF SINGAPORE

2011

**A CORPUS-BASED CONTRASTIVE ANALYSIS OF
SADNESS EXPRESSIONS IN ENGLISH AND CHINESE**

ZHANG RUIHUA
(M. Eng., Shanxi)

**A THESIS SUBMITTED
FOR THE DEGREE OF DOCTOR OF
PHILOSOPHY**

**DEPARTMENT OF ENGLISH LANGUAGE AND
LITERATURE**

NATIONAL UNIVERSITY OF SINGAPORE

2011

ACKNOWLEDGEMENTS

Firstly, I wish to express my deepest appreciation and thanks to my supervisor, Associate Professor Vincent B. Y. Ooi, for his guidance, advice and encouragement throughout the writing of this thesis. As an experienced PhD supervisor, his comments and suggestions are always constructive and it often takes me a while to digest them. From the discussions and meetings with him, I have obtained not only the solutions to specific problems in my thesis, but also insights into how to do research more rigorously. Secondly, I would like to thank Professor Teubert, who inspired me to work on the present topic and offered me a lot of help during my writing. Thirdly, my thanks should go to my husband and daughter, who firmly supported me throughout my PhD study. Fourthly, my sincere thanks go to Professor Bao Zhiming and Dr. Peter K. W. Tan for their critical but highly valuable comments on my thesis and Dr. Joseph Sung-Yul Park for his insightful and helpful remarks on my ISM (Independent Study Module) paper. Fifthly, I am deeply grateful to Associate Professor Chang Baobao, for his permission to use the Babel English-Chinese Parallel Corpus for my research. Finally, my debt of gratitude goes to the National University of Singapore, which offered me the chance and scholarship to complete my PhD project, the staff members of the Department of English Language and Literature and the Division of Research and Graduate Studies, Faculty of Arts & Social Sciences for their full support, and the Central Library and the Chinese Library at NUS for their massive academic resources.

TABLE OF CONTENTS

INTRODUCTION	1
1.1 IMPETUS AND RATIONALE.....	1
1.2 AIMS AND OBJECTIVES	5
1.3 METHODOLOGY.....	8
1.4 SIGNIFICANCE	9
1.5 SUMMARY OF THE CHAPTERS	11
CHAPTER 2.....	12
LITERATURE REVIEW	12
2.1 INTRODUCTION.....	12
2.2 EMOTIONS	12
2.2.1 <i>Emotions in psychology</i>	12
2.2.2 <i>Emotion in linguistics</i>	17
2.3 CONTRASTIVE LEXICAL SEMANTICS.....	27
2.3.1 <i>Lexicalisation</i>	27
2.3.2 <i>The problem of decomposition</i>	29
2.4 CORPUS LINGUISTICS, PARALLEL CORPORA AND CROSS-LINGUISTIC RESEARCH.....	33
2.4.1 <i>Corpus linguistics</i>	33
2.4.2. <i>Translation equivalence, parallel corpora and cross-linguistic research</i>	35
2.5 SUMMARY	40
CHAPTER 3.....	43
THEORETICAL FRAMEWORK	43
3.1 CORPUS LINGUISTICS: THEORY OR METHODOLOGY?	43
3.2 THE CORPUS-LINGUISTIC VIEW ON MEANING	46
3.2.1 <i>Meaning as a social construct</i>	46
3.2.2 <i>Meaning as paraphrase</i>	50
3.2.3 <i>Meaning as usage</i>	52
3.3 A LEXICAL MODEL: COLLIGATION+COLLOCATION+SEMANTIC ASSOCIATION/PREFERENCE.....	53
3.3.1 <i>Sinclair’s notion of (extended) unit of meaning & his lexical model</i>	53
3.3.2 <i>Hoey’s notions of semantic association and lexical priming</i>	57
3.4 SUMMARY	61
CHAPTER 4.....	62
METHODOLOGY	62
4.1 INTRODUCTION.....	62
4.2 CORPORA USED	62
4.2.1 <i>The Bank of English</i>	62
4.2.2 <i>The CCL Corpus</i>	63
4.2.2.1 <i>Comparability</i>	64
4.2.2.2 <i>Representativeness</i>	64
4.2.3 <i>The Babel English-Chinese Parallel Corpus of Fiction</i>	66
4.3 DICTIONARIES USED	66
4.4 SOFTWARE	67
4.4.1 <i>The WordSmith Tools</i>	67
4.4.2 <i>ParaConc</i>	69
4.5 DATA PROCESSING AND ANALYSIS.....	74
4.5.1 <i>Clean-text policy and annotation</i>	74
4.5.2 <i>Data processing</i>	75

4.5.2.1	English data	75
4.5.2.2	Chinese data.....	77
4.5.2.2.1	Chinese data processing	77
4.5.2.2.2	Chinese part-of-speech classification and tagging.....	78
4.5.3	<i>Data analysis</i>	85
CHAPTER 5.....		88
CONTRASTIVE ANALYSIS OF SADNESS EXPRESSIONS IN ENGLISH AND CHINESE		88
5.1	OVERVIEW	88
5.2	CONTRASTIVE ANALYSIS OF SADNESS EXPRESSIONS IN ENGLISH AND CHINESE	89
5.2.1	<i>Sorrow* & grief* and their Chinese equivalents</i>	89
5.2.1.1	<i>Sorrow* and grief*</i>	89
5.2.1.2	Chinese equivalents of <i>sorrow*</i> and <i>grief*</i>	103
5.2.1.2.1	<i>Beitong</i>	105
5.2.1.2.2	<i>Beiai</i>	108
5.2.1.2.3	<i>Beishang</i>	112
5.2.1.3	Comparison.....	115
5.2.2	<i>Unhappy* and their Chinese equivalents</i>	117
5.2.2.1	<i>Unhappy*</i>	117
5.2.2.2	Chinese equivalents of <i>unhappy</i>	120
5.2.2.2.1	<i>Bu gaoxing</i>	122
5.2.2.2.2	<i>Bu yukuai</i>	125
5.2.2.2.3	<i>Bu kuaile</i>	128
5.2.2.3	Comparison.....	130
5.2.3	<i>Sad* and their Chinese equivalents</i>	132
5.2.3.1	<i>Sad*</i>	132
5.2.3.2	Chinese equivalents of <i>sad *</i>	137
5.2.3.2.1	<i>Shangxin</i>	141
5.2.3.2.1	<i>Nanguo</i>	143
5.2.3.2.3	<i>Youshang</i>	146
5.2.3.3	Comparison.....	148
5.2.4	<i>Heartbreak* and their Chinese equivalents</i>	150
5.2.4.1	<i>Heartbreak*</i>	150
5.2.4.2	<i>Xinsui</i>	153
5.2.4.3	Comparison.....	157
5.2.5	<i>Doleful* & woe* and their Chinese equivalents</i>	158
5.2.5.1	<i>Doleful*</i>	158
5.2.5.2	<i>Woe*</i>	161
5.2.5.3	Comparison.....	167
5.2.6	<i>Mourn* and their Chinese equivalents</i>	168
5.2.6.1	<i>Mourn*</i>	168
5.2.6.2	<i>Aidao</i>	173
5.2.6.3	Comparison.....	175
5.2.7	<i>Depression* & Melancholy* and their Chinese equivalents</i>	176
5.2.7.1	<i>Depression* & Melancholy*</i>	176
5.2.7.1.1	<i>Depression*</i>	176
5.2.7.1.2	<i>Melancholy*</i>	181
5.2.7.2	Chinese equivalents of <i>depression* & melancholy*</i>	187
5.2.7.2.1	<i>Jusang</i>	187
5.2.7.2.2	<i>Yayi</i>	188
5.2.7.2.3	<i>Qingxu diluo</i>	190
5.2.7.2.4	<i>Yiyu</i>	190
5.2.7.2.5	<i>Youyu</i>	192

5.2.7.2.6	<i>Shanggan</i>	196
5.2.7.3	Comparison	197
5.2.8	<i>Dejected*</i> , <i>disheartened*</i> & <i>despondent*</i> and their Chinese equivalents	198
5.2.8.1	<i>Dejected*</i> , <i>disheartened*</i> & <i>despondent*</i>	199
5.2.8.1.1	<i>Dejected*</i>	199
5.2.8.1.2	<i>Disheartened*</i>	201
5.2.8.1.3	<i>Despondent*</i>	202
5.2.8.2	Chinese equivalents of <i>dejected*</i> , <i>disheartened*</i> & <i>despondent*</i>	205
5.2.8.2.1	<i>Sangqi</i>	206
5.2.8.2.2	<i>Chuitousangqi</i>	207
5.2.8.2.3	<i>Huixin</i>	207
5.2.8.2.4	<i>Huixinsangqi</i>	208
5.2.8.2.5	<i>Xinhuiyileng</i>	209
5.2.8.3	Comparison	209
5.2.9	<i>Gloom*</i> and their Chinese equivalents	211
5.2.9.1	<i>Gloom*</i>	211
5.2.9.2	Chinese equivalents of <i>gloom*</i>	214
5.2.9.2.1	<i>Yinyu</i>	215
5.2.9.2.2	<i>Menmenbule</i>	216
5.2.9.3	Comparison	218
5.2.10	<i>Upset*</i> and their Chinese equivalents	219
5.2.10.1	<i>Upset*</i>	219
5.2.10.2	Chinese equivalents of <i>upset*</i>	223
5.2.11	<i>Bitter*</i> , <i>agony*</i> , <i>anguish*</i> and <i>tongku</i>	226
5.2.11.1	<i>Bitter*</i>	226
5.2.11.2	<i>Agony*</i> & <i>anguish*</i>	232
5.2.11.3	<i>Tongku</i>	239
5.2.11.4	Comparison	243
5.3	METAPHORS IN SADNESS EXPRESSIONS IN CHINESE AND ENGLISH	246
5.4	A SUMMARY OF THE CORPUS-BASED CONTRASTIVE ANALYSIS	256
CHAPTER 6		264
IMPLICATIONS		264
6.1	INTRODUCTION	264
6.2	IMPLICATIONS FOR RELATED RESEARCH AREAS	264
6.2.1	<i>Psychology</i>	264
6.2.2	<i>Contrastive lexical studies</i>	265
6.2.3	<i>Bilingual lexicography</i>	266
6.2.4	<i>Language pedagogy</i>	267
CHAPTER 7		269
CONCLUSIONS		269
7.1	SUMMING UP	269
7.2	LIMITATIONS	271
7.3	FURTHER RESEARCH	273
BIBLIOGRAPHY		275
APPENDIX: ENGLISH GLOSSING OF CHINESE WORDS		286

SUMMARY

Previous studies of cross-linguistic lexical semantics usually fall within the scope of traditional contrastive lexical semantics, cognitive linguistics, cultural linguistics or Natural Semantic Metalanguage. This thesis attempts to show how contrastive lexical analyses can be carried out by applying a corpus linguistic framework. It adopts a unique theoretical approach which integrates corpus-linguistic theories on meaning (as a social construct, usage and paraphrase) at the macro level with lexical items at the micro level. It develops a new complex methodology which combines computational tools with manual examination to tease meaning out of corpus evidence, to compare and contrast emotion (sadness) expressions that overlap only partly between languages, in this case English and Chinese. Specifically, this research looks at sadness expressions in terms of three corpus-linguistic structural categories, i.e. colligation, collocation and semantic association/preference, and paraphrase (both explicit and implicit) to uncover the syntagmatic lexical relations between these expressions and their syntactically related words and disclose their meanings. The corpus data are from three resources: (1) a monolingual general Chinese corpus, (2) a monolingual general English corpus, and (3) an English-Chinese parallel corpus. By examining the semantic features of semantically-similar emotion expressions on the basis of empirical data across languages, their subtle nuances of meaning can be accurately captured, the culture-specific conceptualisations encoded in their emotion lexicons can be disclosed and their respective cultural distinctiveness of emotion can be highlighted. In addition, this approach can shed insightful light on psychology, bilingual lexicography and language pedagogy.

LIST OF TABLES

TABLE 5.1. 1 THE COMMON PATTERN OF <i>SORROW</i> AND <i>GRIEF</i> IN THE BOE: AND/OR.....	92
TABLE 5.1. 2 CHINESE EQUIVALENTS FOR <i>SORROW*</i> AND <i>GRIEF*</i> IN THE BPC.....	103
TABLE 5.1. 3 DEFINITIONS OF <i>BEISHANG</i> , <i>BEITONG</i> AND <i>BEIAI</i> IN XHCBL.....	103
TABLE 5.1. 4 SIGNIFICANT COLLOCATES OF <i>BEITONG</i> IN THE CCL CORPUS	105
TABLE 5.1. 5 SIGNIFICANT COLLOCATES OF <i>BEIAI</i> IN THE CCL CORPUS	108
TABLE 5.1. 6 SIGNIFICANT COLLOCATES OF <i>BEISHANG</i>	112
TABLE 5.1. 7 CONTRASTIVE ANALYSIS: <i>SORROW & GRIEF</i> VS <i>BEITONG</i> , <i>BEISHANG & BEIAI</i>	116
TABLE 5.2. 1 THE MOST FREQUENT CHINESE EQUIVALENTS OF <i>UNHAPPY*</i> IN THE BPC.....	120
TABLE 5.2. 2 SIGNIFICANT COLLOCATES OF <i>BU GAOXING</i> IN THE CCL CORPUS.....	122
TABLE 5.2. 3 SIGNIFICANT COLLOCATES OF <i>BU YUKUAI</i> IN THE CCL CORPUS	126
TABLE 5.2. 4 SIGNIFICANT COLLOCATES OF <i>BU KUAILE</i> IN THE CCL CORPUS	128
TABLE 5.2. 5 COMPARISON OF SOME COLLOCATES OF <i>BU GAOXING</i> , <i>BU YUKUAI & BU KUAILE</i>	131
TABLE 5.2. 6 FEATURES OF UNHAPPY EXPRESSIONS.....	131
TABLE 5.3. 1 THE MOST FREQUENT CHINESE EQUIVALENTS OF <i>SAD*</i> IN THE BPC	138
TABLE 5.3. 2 SIGNIFICANT COLLOCATES OF <i>SHANGXIN</i> IN THE CCL CORPUS	141
TABLE 5.3. 3 SIGNIFICANT COLLOCATES OF <i>NANGUO</i> IN THE CCL CORPUS	144
TABLE 5.3. 4 COMPARISON OF THE COLLOCATES FOR <i>SHANGXIN & NANGUO</i>	145
TABLE 5.3. 5 SIGNIFICANT COLLOCATES OF <i>YOUSHANG</i> IN THE CCL CORPUS.....	147
TABLE 5.4. 1 CHINESE EQUIVALENTS OF HEARTBREAK* IN THE BPC	153
TABLE 5.4. 2 SIGNIFICANT COLLOCATES OF <i>XINSUI</i> IN THE CCL CORPUS	155
TABLE 5.5. 1 CHINESE EQUIVALENTS OF <i>WOE*</i> IN THE BPC	164
TABLE 5.6. 1 CHINESE EQUIVALENTS OF MOURN* IN THE BPC	173
TABLE 5.6. 2 SIGNIFICANT COLLOCATES OF <i>AIDAO</i> IN THE CCL CORPUS.....	174
TABLE 5.7. 1 AND/OR PATTERN FOR <i>DEPRESSION</i> IN THE BOE.....	179
TABLE 5.7. 2 TOP 3 CHINESE EQUIVALENTS OF <i>DEPRESSION*</i> & <i>MELANCHOLY*</i> IN THE BPC.....	187
TABLE 5.7. 3 SIGNIFICANT COLLOCATES OF <i>JUSANG</i> IN THE CCL CORPUS	187
TABLE 5.7. 4 SIGNIFICANT COLLOCATES OF <i>YAYI</i> IN THE CCL CORPUS	189
TABLE 5.7. 5 SIGNIFICANT COLLOCATES OF <i>YIYU</i> IN THE CCL CORPUS	191
TABLE 5.7. 6 SIGNIFICANT COLLOCATES OF <i>YOUYU</i> IN THE CCL CORPUS	192
TABLE 5.7. 7 SIGNIFICANT COLLOCATES OF <i>SHANGGAN</i> IN THE CCL CORPUS	196
TABLE 5.8. 1 SIGNIFICANT COLLOCATES OF <i>HUIXIN</i> IN THE CCL CORPUS.....	207
TABLE 5.8. 2 COMPARISON OF <i>DEJECTED*</i> , <i>DESPONDENT*</i> AND <i>DISHEARTENED*</i> AND THEIR CHINESE EQUIVALENTS	209
TABLE 5.9. 1 CHINESE EQUIVALENTS OF <i>GLOOM*</i> IN THE BPC	214
TABLE 5.9. 2 SIGNIFICANT COLLOCATES OF <i>YINYU</i> IN THE CCL CORPUS	215
TABLE 5.9. 3 SIGNIFICANT COLLOCATES OF <i>MENMENBULE</i> IN THE CCL CORPUS	217
TABLE 5.11.1 CHINESE EQUIVALENTS OF <i>BITTER*</i> IN THE BPC.....	230
TABLE 5.11. 2 CHINESE EQUIVALENTS OF <i>AGONY*</i> IN THE BPC	238
TABLE 5.11. 3 CHINESE EQUIVALENTS OF <i>ANGUISH*</i> IN THE BPC.....	238
TABLE 5.11. 4 SIGNIFICANT COLLOCATES OF <i>TONGKU</i> IN THE CCL CORPUS.....	241
TABLE 5.13.1 COMPARISON OF METAPHORS IN SADNESS EXPRESSIONS IN CHINESE AND ENGLISH	247
TABLE 5.13. 2 MAJOR DIFFERENCES BETWEEN ENGLISH AND CHINESE SADNESS EXPRESSIONS.....	259

LIST OF FIGURES

FIGURE 4. 1 A SCREENSHOT OF THE COLLOCATES OF 伤心 BY WORDSMITH 5.0	68
FIGURE 4. 2 A SCREENSHOT OF SEARCHING RESULT BY PARACONC	72
FIGURE 4. 3 THE INTERFACE OF THE CONCORDANCING OPTIONS OF WORD SKETCH.....	76
FIGURE 5. 1 CONCORDANCE LISTING OF <i>MINZU DE BEIAI</i>	109
FIGURE 5. 2 DIAGRAM FOR CONTRASTIVE CORPUS-BASED ANALYSIS	257
FIGURE 5. 3 EQUIVALENCE NETWORK OF SADNESS EXPRESSIONS BETWEEN ENGLISH AND CHINESE ..	258

LIST OF ABBREVIATIONS

BOE	The Bank of English
BPC	The Babel English-Chinese Parallel Corpus of Fiction
CCL	The Chinese corpus compiled by the Centre for Chinese Linguistics, Peking University, China
CDA	Critical Discourse Analysis
COBUILD	Collins COBUILD Advanced Learner's English Dictionary
LDCE	Longman Dictionary of Contemporary English
OALECD	Oxford Advanced Learner's English -Chinese Dictionary
OED	Oxford English Dictionary
POS	part of speech
XHCBL	Xiandai Hanyu Cidian: Bilingual Edition (A Dictionary of Modern Chinese)
5 th XHC	Xiandai Hanyu Cidian (A Dictionary of Modern Chinese) 5 th Edition

CHAPTER 1

INTRODUCTION

1.1 Impetus and Rationale

The issues of universality versus cultural specificity of emotions have long been debated among psychologists and biologists. Biologists tend to believe that emotions are innate and universal because they have biological bases. The dominant view among psychologists, mainly based on some psychological experiments, is that human beings have some basic emotions, which are universal. The cultural specificity of emotions cannot be discussed without recourse to the cultural specificity of language, which has often been overlooked by psychologists. To what extent does language shape the way we perceive things and therefore influence our emotions? Whorf (1956) tells us that the way we cut the world into objects and actions is not a true reflection of reality but a division imposed upon us by our language. Although the strong version of the Sapir-Whorf hypothesis---language entirely determines thought and cognition--is probably not tenable, more and more linguists have begun to support its weak version, i.e. language does influence our thoughts to some extent. According to this hypothesis, the emotional lexicon of a language constrains the thoughts of its speakers about feelings. Each language has its own set of emotion words to characterise emotional experience, but apparently there are more distinct human emotions than those words can describe (Wierzbicka, 1992). A language has not lexicalised a particular emotion because its culture has not found it worthy of a specific word; it does not mean that its speakers

have never experienced that emotion or cannot perceive it as a recognizable feeling (Wierzbicka). This is why we reject the Sapir-Whorf hypothesis in its strong form. However, the embracing of its weak version means that the ways in which the speakers of a language characterise, discuss and communicate their emotional experience are largely shaped by the ready-made emotional lexicon in their language.

‘Emotion words reflect, and pass on, certain cultural models, and these models, in turn, reflect and pass on values, preoccupations, and frames of reference of the society (speech community) within which they have evolved’ (Wierzbicka, 1999: 32). The close relation between language and culture can be explained by the newly rising enterprise--cognitive linguistics¹, represented by the works of Lakoff & Johnson (1980), Lakoff (1987), Langacker (1987) and the like. Its basic assumption is that human language cannot be viewed as an abstract system isolated from socio-psychological settings in which it is used; on the contrary, it reflects the way people experience the world. In other words, a wide variety of factors such as perception, reasoning, the nature of the body, the emotions, memory, social structure, and cognitive development etc. determine the structural characteristics of language. Langacker (1999: 16) reveals the correlation between language and culture:

...the specific conventions of a given language are culturally transmitted through social interaction. Language is an essential instrument and component of culture, whose reflection in linguistic structure is pervasive and quite significant.

¹ Cognitive linguistics in the thesis does not refer to Chomskyan linguistics, but a branch of linguistics which argues that language is only a part of human cognition, represented by Lakoff, Langacker, Fauconnier et al.

At the outset of communication, speakers of a language need to characterise the objects and events in their world. If we reject the innate hypothesis, their characterisations are unavoidably influenced by the way in which they perceive and construe things around them. Such characterisations are then encoded into various layers of their language and culturally transmitted from generation to generation; of course, they have been gradually enriched in the later development of the language. Hence, the organisation of a language is shaped by the way in which its speakers perceive the world and conceptualise the phenomena around them. This explains, in large measure, why different languages have different structures and lexicons. The lexicalisation of a concept is to 'present it as an established category of human thought' (Fillmore, 2003: 258). As the established categories of human thought, the lexicon of a language will no doubt reflect its cultural features, including its emotional lexicon. Semantically similar emotion words in different languages might overlap. The comparison of the emotion lexicon of a language with that of another can therefore suggest similarities and differences in conceptualisations of emotion.

In lexical contrastive semantics, one of the thorniest problems is how to establish equivalence of lexical items across languages. One common belief would be that equivalent lexical items in different languages can be decomposed into identical set of features. This sort of decomposition presupposes the existence of primitive or universal concepts, which can serve as *tertium comparationis* to establish lexical equivalence across languages. However, this will bring numerous problems. Krzeszowski (1990) presents a detailed analysis of this difficulty (see Section 2.3.1). It

seems that the establishment of equivalence of lexical items across languages, by decomposing words into primitive concepts, is more than difficult to achieve. This is because the idea of primitive concept itself is problematic: we have to express it in a natural language, where the great majority of words are polysemous and ambiguous. Traditionally, their meanings in the dictionary have to rely on other words there which in turn will unavoidably lead to circularity. Meaning is far more subtle than a componential analysis based on discrete features can account for, and far more complex than any print dictionary can describe. To carry out a cross-linguistic lexical study, the better choice is to start with 'a perceived or assumed similarity' between contrasted items (James, 1980: 168) rather than to seek primitive or atomic units. What we can do is to refine the 'initial assumptions of similarity' (Altenberg & Granger, 2002: 16). The analysis is only to 'add explicitness, precision, perhaps formalization' (Chesterman, 1998: 58); it may also provide 'added information, added insights, added perception' (Chesterman).

Most of the traditional contrastive lexical studies are incomplete and inadequate, for they are not supported by empirical quantitative data and fail to present linguistic information on frequency of use, collocations, colligations, semantic preferences/associations of the examined items. Their lexical descriptions have been largely based on individual researcher's introspection, dictionaries or scanty empirical evidence. Because of the inadequacy of previous contrastive lexical studies, there has been a growing tendency to base such studies on corpora. The methodology of corpus linguistics is based on scientific principles. It relies on authentic texts and computational

tools to detect and classify linguistic phenomena; the datasets and computer programs are open to scrutiny by the peer community; the results they produce are reproducible. This approach to lexical studies aims to arrive at the meaning of a word by observing its external linguistic behaviour in various contexts rather than looking for its primitive units intuitively. Such a practice can be justified by Firth's (1957) dictum that a word shall be known by the company it keeps. In keeping with Firth's observation, Levin's (1993: 1) central claim that the behaviour of a verb is to a large extent determined by its meaning can also be extended to other classes. Thus, the semantic information encoded in the word can be revealed by examining its linguistic behaviour. By using corpus tools, we can read its concordance and examine its collocates. As long as we have sufficient instances, anything significant about the node word will be disclosed. Hence, to know the meaning of a word does not necessarily follow that we have to decompose it. Computer corpus linguistics will tell us in exactly what way collocation is related to meaning. Corpus-assisted contrastive lexical studies will show that distinctions of semantically similar words in the same language or in different languages can be drawn without decomposing them.

1.2 Aims and objectives

Emotions are not innate but constructed in various cultural contexts through the process of socialization (e.g. Santangelo, 2003). Different cultural groups conceptualise their emotional experience in different ways, as they categorise other human actions and thoughts (e.g. D'Andrade, 1995; Palmer, 1996; Strauss & Quinn, 1996). The

conceptualisations emerging at the cultural level of cognition are referred to as ‘cultural conceptualisations’ (Sharifian, 2003; 2008), which are negotiated and renegotiated by its members through time and across generations (Sharifian). The cultural conceptualisations of emotions are encoded in the emotional lexicon of a language, which reflect the ideas, beliefs and cultural models of emotions of its speech community. Analysis of the emotional lexicon of a language helps reveal the concepts available to its speakers through which they categorise and describe their emotional experience. Where are lexicalised in a single word, the presence of a word implies the existence of the concept it expresses. Language-specific emotion words might overlap semantically across languages. By comparing and contrasting the semantic features of the emotion words in a language and their equivalents in another, we can know what they share and how they differ.

Sadness has been proposed as one of the basic emotions by many psychologists. However, what are considered basic emotions in the Chinese culture are markedly different from those proposed for the English language. The disparity between sadness concepts across these two cultures seems to be more significant, as compared with happiness, anger etc. The primary goal of this research is to investigate the general concept of ‘sadness’ within the English and Chinese contexts by examining lexical sets of sadness expressions. The exploration will be conducted by applying corpus-linguistic theories (Teubert, 2005, 2010; Teubert & Cermakova, 2004) on meaning (as a social construct, usage and paraphrase), i.e. meaning is language use and it is constructed and negotiable through paraphrases in discourse, and by employing Hoey’s

(2005) or adapted Sinclair's (1996; 2004) corpus-linguistic lexical model, i.e. to look at lexical items in terms of colligations, collocations and semantic preferences/associations.

In my thesis, I will describe the colligational properties of sadness expressions and formulate their semantic preferences/associations, compare and contrast the semantically similar items within and across languages to capture their semantic similarities and differences, uncover the respective cultural distinctiveness, and explain why they exhibit such differences from a cultural perspective. Their semantic information will become vivid when we look at the empirical language data in corpora. As long as our corpora are large enough, our quantitative tools will yield sufficient data, on the basis of which the exact meanings of these sadness expressions and the finer distinctions between them can be accurately captured. By comparing the semantic features of sadness expressions between Chinese and English, their emotive content will be revealed and the cultural difference of emotion between English and Chinese can be highlighted.

The sadness expressions to be examined include (cf. Section 5.1): *sorrow, grief, unhappy, sad, heartbreak, mourn, doleful, woeful, woesgone, depression, melancholy, gloomy, upset, dejected, disheartened, despondent, bitter, anguish, agony*, in English and *beitong, beishang, beiai, shangxin, nanguo, youshang, bu gaoxing, bu kuaile, bu yukuai, xinsui, aidao, choumeikulian, jusang, yayi, qingxudiluo, yiyu, youyu, shanggan, yinyu, menmenbule, huixin, sangqi, chuitousangqi, huixinsangqi, xinhuiyileng, tongku* in Chinese. The above English items will be looked at together with their cognate items. For instance, *sad* will be examined along with *sadness, sadden, sadly*. Finally, I will

present an equivalence network between English and Chinese sadness expressions and a summary of their semantic similarities and differences across languages.

1.3 Methodology

There has been an increasing awareness that cross-linguistic studies cannot rely on introspection or dictionaries, but must be firmly based on naturally occurring language used in a variety of situations. With the aid of computer technology, we can observe semantic similarities and differences across languages objectively; we can identify clearly what elements they share with similar expressions from the other language and what elements are specific to their own language. Basing my research on corpora ensures that my investigation leads to tenable results and a more accurate description of English and Chinese sadness expressions.

Discussions of emotions can be cleared from the charge of being impressionistic by means of an examination of empirical data. To retrieve reliable information on lexical use, large corpora are needed especially for less frequent words. Sufficient instances of the examined item from various texts are required to give a reliable picture about its use. The data resources of my investigation will be the Chinese corpus compiled by the Centre for Chinese Linguistics, Peking University, the Bank of English and the Babel English-Chinese Parallel Corpus of Fiction compiled by Peking University. This corpus-based approach is embedded in a corpus-linguistic theoretical framework. The purpose of incorporating a parallel corpus into this methodology is to base cross-linguistic lexical links on quantitative data.

This analysis will involve the colligations and semantic preferences/associations of the English and Chinese sadness expressions by examining their collocates. Colligations ‘provide an objective view from the outside. They describe the observable behaviour of a social group’ (Stubbs, 2009: 125). According to Stubbs, semantic preferences/associations ‘provide a subjective view from the inside’ and they can only be identified intuitively (p.125). Sometimes it is not an easy task to label them adequately. Once this is achieved, I will then look for real paraphrases (explicit paraphrases) and typical examples (implicit paraphrases, see Section 3.2.2) that contain information contributing significantly to revealing meaning. The final step of the analysis is to compare and contrast their meanings on the basis of the previous findings and provide a cultural explanation for the differences.

1.4 Significance

Psychological approaches to emotions have led to a number of important insights. However, such studies have the disadvantage of seemingly less objective because they are either based on introspection or the subjects studied are unlikely to behave in a natural way under observation. On the other hand, the corpus-based approach to the study of emotions offers a number of advantages. Firstly, there is no ‘observer’s paradox’; in other words, since the study is based on texts written independently, the results will be less distorted by the subjects. Secondly, it is easier to achieve accuracy and reliability in such an analysis because the number of texts to be examined is much larger than can be looked at in psychological studies.

This study extends the corpus-linguistic framework to analyse and compare lexical meaning in depth. It proposes a unique theoretical approach which integrates corpus-linguistic theories on meaning (as a social construct, usage and paraphrase) at the macro level with lexical items at the micro level. It also develops a new complex methodology, which combines computational tools with manual examination to tease meaning out of corpus evidence, to compare and contrast lexical meaning within and across languages. Hopefully, this will provide fresh insights into how developments in corpus-linguistic theories and in the application of computer corpus technology can advance the field of lexical contrastive linguistics. By so doing, we can have a nuanced description of similarities and differences that epitomises the descriptive potential of the corpus-linguistic approach to contrastive lexical semantics. My thesis will be the first large-scale corpus-based examination of the ‘sadness’ expressions. By analysing a large quantity of contrastive corpus data, I will uncover differences in the use of emotion words, and a large part of my thesis will be devoted to explaining them. I will show to what extent the differences can be explained through cultural differences. This is a relatively recent field of enquiry and investigations are still few and far between. My own contribution will hopefully offer new insights not yet described in the literature. In addition, the corpus-linguistic approach can throw new light on psychology, contrastive lexical studies, bilingual lexicography and language pedagogy.

1.5 Summary of the chapters

This thesis explores, with reference to large corpora, the semantic disparities between sadness expressions in English and Chinese by applying the corpus-linguistic approach. It aims to probe into its cultural reasons which caused these differences, based on the corpus findings. Chapter 1 gives the rationale, objectives, methodology of this research, and its significance. Chapter 2 reviews the relevant literature on emotions in psychology and linguistics, contrastive lexical studies and corpus linguistics. Chapter 3 explicates the theoretical framework on which the research is based and Chapter 4 deals with its methodology. Chapter 5 is the body of the thesis, which mainly reports, with great detail, on the corpus findings from the semantic-contrastive analysis of English and Chinese sadness expressions. The corpus-based analysis is also complemented by a contrastive-qualitative analysis of metaphorical sadness expressions in English and Chinese. Chapter 6 addresses possible implications of this research for psychology, contrastive lexical semantics, bilingual lexicography and language pedagogy. The final chapter summarises the research findings, discusses the limitations of this research and envisages the further development of cross-linguistic lexical studies with the aid of the corpus-linguistic approach in the near future.

CHAPTER 2

LITERATURE REVIEW

2.1 Introduction

This research is grounded in three research areas: emotions, contrastive lexical semantics and corpus linguistics. Section 2.2 addresses topics of emotions in psychology and linguistics. Contrastive lexical semantics is dealt with in Section 2.3, which mainly discusses the frequently-discussed issues of lexicalisation and decomposition. Section 2.4 is concerned with corpus linguistics, parallel corpora and cross-linguistic research. The last section concludes that it is time to embrace corpora in studies of emotions and contrastive lexical semantics.

2.2 Emotions

2.2.1 Emotions in psychology

There has been heated debate about whether emotions are cultural specific or universal among anthropologists, biologists and psychologists. In 1872, Darwin addressed the behavioural manifestations of emotions in his *The Expression of the Emotions in Man and Animals* (1872/1965). Much of the research on emotion in the field of psychology maintains his tradition by documenting universal aspects of human emotions. The basic emotion theory has drawn the most attention, which argues that there are 5-7 basic emotions common to all human beings. Proponents of this theory have used facial-expressions recognition studies as their central source of evidence. Their assumption is that facial expressions are the translations of psychological states or processes. For

example, Izard (1971) shows that facial expressions of Americans in photographs expressing 'basic' emotions can be recognised above chance level in both preliterate and literate cultures and she concludes that these facial expressions of emotions are universally recognised. Ekman (2007) carries out similar experiments over many decades and draws a similar conclusion, namely that basic emotions are inborn and everyone understands them. In these experiments, people in different countries---Papua New Guinea, the United States, Japan, Brazil, Argentina, Indonesia and the former Soviet Union, listen to stories and later are asked to match these stories with one of the photographs of facial features expressing seven basic emotions (sadness, anger, surprise, fear, enjoyment, disgust, and contempt). These studies which aim to demonstrate universality have been criticised on the grounds that their methodology biases their outcome. In particular, it seems unreliable to tell participants to view expressions and enter multiple choice responses from a predetermined list. As Russell (1991) argues, what such experiments can show is that people from different cultures give similar (but not necessarily identical) interpretations to facial expressions. He says, suppose you are asked to match one of the terms to a face wearing a bright smile:

Most likely, you'd select happy. But now suppose that happy had been replaced on the list with elated. Given the alternatives, you'd have no choice but to select elated. If happy were successively replaced with serene, satisfied, excited, grateful, and triumphant, you'd again select any of these words in turn....Indeed, substitute for happy any clearly positive word... and the conclusion remains the same (p. 435).

Moreover, it remains unclear how translation equivalence is assured. For instance, in the experiment of happiness in Papua New Guinea, Ekman uses 'His /her friends have come and s/he feels very happy' for the examinees to match with photographs of

emotions (2007). However, after being translated from English to Pidgin and then to Fore, the stories are told to the examinees finally in Fore, a local language in Papua New Guinea. Due to the great semantic variations between languages, it is problematic to assume, on the basis of such experiments, that the people in those remote and isolated areas understand emotions in the same way that western people do and claim further that all these emotions are universal and have a genetic foundation.

Furthermore, often the same studies demonstrating universality also implies cultural specificity of emotions (Elfenbein & Ambady, 2002). In Ekman et al's (1969) study, the recognition accuracy of the pre-literate participants from New Guinea, albeit above the chance level, is apparently less than that achieved by the literate participants, showing that there is still cultural variation in the accuracy of recognition. The cultural differences are also marked in Izard's (1971) study, in which American & European, Japanese and African groups identify 75-83%, 65% and 50% of the facial photographs respectively. Matsumoto (1989) argues that collectivistic cultural groups are less accurate in expressing and perceiving negative emotions because of the risk of disrupting social disorder. Emotional communication may be more accurate when the expresser and the perceiver are both from the same cultural group rather than from different cultural groups (Elfenbein & Ambady, 2002).

On the other hand, cultural constructionists claim that emotions are culturally organised or moulded, and thus are culture-specific. Some early ethnographic work on emotion have showed convincingly that cultures vary in which emotions are aroused in a given situation and how particular emotions are conceptualised and labelled (Briggs,

1970; Lutz, 1988). Durkheim (1911/1961) terms 'collective conscience' the ideas, norms, and social expectations held important in the minds of all members of a society and points out that feelings are closely linked to social situations. Kleinman (1980) suggests affect comes into consciousness when a state of psychological arousal is given shape by a cultural judgment, the process of assigning an emotion label, made by individual who suffers the arousal. Some recent work in the anthropology of emotion has also critiqued the notion that human emotional expressions are universal and have their basis in biology (Shweder, 1994) by investigating affect as a cultural category and symbolic system. Secondary or cognised affects in different cultures, once labelled, differ in quality or intensity. Culture provides the cognitive and social environment to which the individual must learn to accommodate.

Goddard (1997) argues that psychology has yet to fully come to grips with the problem of translation between languages; researchers in cultural psychology and ethnopsychology (e.g. Harre & Finlay-Jones, 1986; Lutz, 1988; Shweder, 1991, 1993; White & Kirkpatrick, 1985) are well aware of the semantic variation between languages, while it is not the case in the field of psychology at large. Goddard takes as an example a recent edited volume entitled *The Nature of Emotion* (Ekman & Davidson, 1994), where the terms 'translation' and 'language' are not included in the index, and only one of over 20 contributors acknowledges that languages issues would pose 'any great conceptual or methodological dilemma' (p.155). Statements like this are more often seen in books of psychology: 'Equivalentents for most of the emotion terms commonly considered part of the "basic" category seem to exist in almost all major languages of

the world' (Scherer, 1994: 25). In some translation-sensitive studies of emotion, the researchers first decide on the emotion terms in English, then translate them into the respective other languages, and finally back-translate them into English. However, it is questionable that the procedure of 'back-translation' can really make sure semantic equivalence across languages. In a series of studies by Jerry Boucher and associates (Boucher & Carlson, 1980; McAndrew, 1986), a pair of Malay expressions (*marah* and *naik darah*) is used as an equivalent of English 'anger'. But Goddard's study (1996) shows that *marah* is closer to English 'offended' than to 'angry', and *naik darah* is more like 'to lose (one's) temper'. If cultures and languages differ in conceptualising and expressing emotions, how can the exact equivalents across languages be found? What they can get are only close or near equivalents available in those languages.

Goddard's contrastive study (1997) of 'surprise-like' words from Malay (*terkejut, terperanjat, hairan*) and English (*surprised, amazed, shocked, startled*) shows that the English word *surprise* does not have an equivalent counterpart in Malay. He finds no semantic core shared by the various 'surprise' expressions either, only 'a loose set of cross-cutting and overlapping semantic correspondences' (p.153). His findings are at odds with the traditional 'basic emotions' theory, which would regard *surprise* as one of the universal human emotions. Actually, it has also been argued that traditional psychological categories, such as 'emotion', 'sensation' and 'cognition', are also cultural artefacts of western society (Harre, 1986; Lutz, 1988). In the same article, Goddard (1997) points out that English should not be used uncritically as the metalanguage of cross-cultural research; English folk taxonomy should not be assumed

as objective psychological reality; indigenous emotions concepts should not be tagged with English glosses. He claims that cross-cultural psychology can ill-afford to ignore the extent of semantic variation between languages; recent developments in linguistic semantics can benefit psychological research by overcoming 'some of the conceptual confusions and methodological difficulties impeding the development of a soundly based cross-cultural psychology' (p. 154).

2.2.2 Emotion in linguistics

To understand emotion is to ask for its origin in development. Psychological research shows that the emotional development of the child is deeply embedded in the emotional culture. From this developmental perspective, the child is 'constantly experiencing complex, constantly changing, internal and external emotional "stimuli", including the continuously changing expression of other people, filtered by cultural rules and norms, and the ongoing flow of language; [i]n these highly complex situations emotional meaning is co-constructed' (Josephs, 1995: 286). This is exactly what cultural constructionists claim: emotions are culturally organised or moulded, and thus are culture-specific. The cultural specificity of emotions cannot be discussed without recourse to the cultural specificity of language, the discussion of which originates from Sapir (1933: 166):

Languages differ widely in the nature of their vocabularies. Distinctions which seem inevitable to us may be ignored in languages which reflect an entirely different type of culture, while these in turn insist on distinctions which are all but unintelligible to us. Such differences of vocabulary go far beyond the names of cultural objects such as arrow point, coat of armour, or gunboat. They apply just as well to the mental world.

Vocabulary is the best evidence of the reality of ‘culture’ because it reveals the specific concepts and classifications of categories and has been historically transmitted. Emotion words in different languages favour different distinctions as well, reflecting their respective cultures. The emotional lexicons of various languages differ significantly, and this reflects considerable differences in ideas and beliefs about emotions and cultural conceptualisations of emotions. These culture-bound emotion words might overlap. And, ‘presumably, the closer two cultures are, the greater the overlap between their respective sets of emotions words’ (Wierzbicka, 1992b: 124). Interpretative categories that have been applied to emotions, such as *surprise* in English, or *jingya* in Chinese, are language- and culture-specific. In turn, the categorisation of feelings depends mainly on ‘the introspective vocabulary of the seeker’ (W. James, 1890: 485). In other words, the emotional lexicon in one’s native language, to a great extent, constrains what kind of feelings people think they might have.

In cross-cultural lexical semantics, the theory that has drawn the most attention is the Natural Semantic Metalanguage (NSM) framework mainly developed by Wierzbicka (1972, 1992b, 1995, 1996, 1999). Within this framework, the meaning of a word is phrased in terms of small, standardised and translatable metalanguage, i.e. ‘universal concepts’ or ‘semantic primitives’ (also called ‘conceptual primitives’ or ‘lexical universals’) such as GOOD, FEEL, LIKE, HERE, THINK, WANT, DO and HAPPEN. These concepts are said to exist in all or most of the world’s natural languages. Wierzbicka argues that the great majority of the words in natural languages are culture-bound, including emotion terms. She believes that each language with its

own lexicon imposes its own classification and interpretative categories upon the world, which vary a great deal across languages and cultures. However, she also believes that there are some semantic components, be it a separate word, an affix or a fixed phrase, which are present in almost all languages, such as PEOPLE, GOOD, DO, or HAPPEN. These universal concepts (about 60 have been identified) can be employed to define or explain other more complicated cultural-specific concepts without the danger of 'terminological ethnocentrism'. The central claim of this theory is that the interpretation of any word (other than the universal concepts) in any language can be reduced to a combination of these semantic primitives. In other words, what Wierzbicka and her followers have been endeavouring to do is to deconstruct the words in natural languages, compare them and find what semantic components in them are common and what different. She states, '[T]he use of conceptual primitives allows us to explore human emotions...from a universal, language-independent perspective' (1995: 236).

Wierzbicka's search for conceptual universals has led to a substantial body of findings on cross-linguistic lexical semantics. Examples examined are from English, Russian, Polish, Ifaluk (Wierzbicka, 1986, 1992a, 1992b, 1995, 1999), Yankunytjatjara and Malay (Goddard, 1990, 1991, 1995, 1996, 1997), Aboriginal English and Maori (Harkins, 1994, 1996) and Chinese (Ye, 2001), among other languages. To give a clear illustration of this approach, three 'anger-like' words from English, Malay and Ifaluk (cf. Goddard, 1996; Lutz, 1987, 1988; Wierzbicka, 1992b: 147) are used to show how their meanings are described by using so-called 'semantic primitives' (adapted from Goddard, 1997: 160-161):

It will be seen that the meanings are depicted by reference to a prototypical cognitive scenario in which a person thinks certain thoughts (often in association with certain desires) which give rise to a pleasant or unpleasant feeling. The cognitive scenarios for these three words all begin with the thought ‘this person did something bad’ and give rise to an unpleasant feeling.

(a) anger (English)

when X thinks of Y, X thinks something like this:

this person did something bad

I don’t want this

I want to do something to this person because of this

because of this, X feels something bad

(b) marah (Malay) =

when X thinks of Y, X thinks something like this:

this person did something bad

this person knows I do not want him/her to do something like this

I feel something bad because of this

I want this person to know this, not because I say anything about it

Because of all this, X feels something very bad

(c) song (Ifaluk)

when X thinks of Y, X thinks something like this:

this person did something bad

I don’t want this

It is bad if people do things like this

I want this person to know this

I want to do something because of this

Because of all this, X feels something bad

We can see that the definitions are ‘stated in the form of an explanatory paraphrase’ (Goddard, 1997: 159), which are made up of ‘semantic primitives’ all languages have in common. According to Wierzbicka and Goddard, the NSM approach can show what the cross-cultural concepts have in common and how they differ. In this example, the three words all denote a bad feeling evoked by the thought that someone has done something bad. But, *anger* is ‘active and focused on retaliation’, and *marah* is ‘muted and restrained’, while *song* is ‘righteous and not necessarily directed toward the offering party’ (Goddard, 1997: 161).

Though the NSM theory's aim to reveal the cultural-specificity by using cultural-free and language-independent concepts is laudable, Bamberg (1997) argues that 'the methodological tool in the form of a universal inventory of lexical items' is unable to perform this job (p.183). He gives the following reasons (p.183):

First, ...the natural semantic metalanguage falls short of delivering more or better insights than traditional interpretative approaches to human actions and texts. Second, the discursive orientations and perspectives within which emotion terms are put to use in actual talk are washed out. Consequently, in order to achieve some richer understanding of emotion terms, particularly those of a foreign language, fuller contours of emotion talk need to be taken into account.

As the quote shows, he thinks the NSM theory is no better than the traditional interpretative approach. Goddard (1997), a keen follower of Wierzbicka, claims that the NSM methodology will lead to a better understanding of word meanings. However, Bamberg casts doubts on it by saying that 'most interpretative ethnographers, including Geertz or Garfinkel (1967), as well as semanticists of the Fillmore/Lakoff tradition, have always stressed that the explication of lexical meanings is never complete, but at best an approximation' (1997: 187). Bamberg's second reason concerns perspectives. The descriptions of the emotion scenarios under the NSM are presented through the shifting of roles from the first-person to a third-person perspective, assuming that it does not make any difference whether the person is the undergoer, instigator or onlooker. Nevertheless, Bamberg's work repeatedly shows that 'perspective matters tremendously in how emotion scenarios are construed and understood' (1997: 187). For example, a presentation of an anger event from the first-person perspective (when I once was angry)

is different from one from a third-person perspective (when John once was angry) both in structure and content.

Josephs (1995) also makes some critical comments on Wierzbicka's approach. He points out that the NSM has brought novel analytic schemata to the semantics of emotions, but it still overlooks a number of relevant issues. Firstly, 'her static and exclusively structural approach to emotion excludes functional aspects of emotions' (p. 279). Emotions are not static but ongoing processes. Also, emotions are 'deeply-rooted in human action and functional in intrapersonal and interpersonal communication' (p. 286). So we cannot cut them off from their context of usage. The meaning of the emotion, i.e. its functional significance, should be found within the socio-cultural system (Averill, 1980). Besides, Wierzbicka's understanding of semantics is problematic in the sense that she analyses the meanings of feelings, isolated from its usage, from the process of encoding and decoding. She assumes that words can be decomposed into cognitive components, 'the sum of which makes the whole—an assumption over which advocates of gestalt psychology especially would stumble' (Josephs, 1995: 282). Furthermore, Wierzbicka's position on universality is questionable. If we believe that lexical universals do exist, then where are they? Are they in dictionaries or people's brains or somewhere else? The descriptions of the cognitive scenarios under her framework are actually highly individual choices, thus having little to do with universality (Josephs).

To sum up, firstly, it remains problematic whether the notion of 'semantic primitives' itself can hold water since they are English words themselves. It is also

unclear whether we can treat them truly neutrally or language independently. Wierzbicka and Goddard would take 'I' as a semantic primitive, but Harre's (1993) investigation shows that 'I' does not mean the same across all languages. Similarly, what is taken as 'good' or 'bad' is not the same across all cultures. For instance, *missing* is a beautiful feeling in the Chinese culture, while it is a negative one in the English-speaking society. So it is unnecessary to approach culture-specific concepts by inventing or discovering lexical universals. The understanding of foreign practices is possible by looking at them from your own perspective and 'not by giving up one's own, adopting a neutral universal one, and from there approaching the foreign' (Bamberg, 1997: 190). Concerning this, Bamberg makes the following comments (1997: 190):

The idea of a full, complete and objective understanding lurking behind the metaphor of semantic universals may even obstruct a better understanding of other cultures and their communications, because it posits the possibility of perfection and truth, while traditional interpretative and hermeneutic approaches are satisfied with the approximations that are desirable and in constant need of being reflected in light of historical changes and cultural affordances. In sum, then, I do not see any reason to give up well-established interpretative practices that have been tested in the interpretation of foreign texts and cultures in exchange for the promises of the natural semantic metalanguage approach.

Besides, Rosch's (1973) position that concepts are not defined by giving a list of features but by describing 'typical' examples is certainly not right in every way, but it is reasonable in some sense.

Secondly, it is highly simplistic to think the description of the world can be broken down to 60 basic universal concepts. It is hard to imagine words with complicated meanings can be defined or explained by these so-called semantic

primitives. Thirdly, Wierzbicka has been trying to capture and depict the isolated meanings of single words without looking at their contexts, frequencies, colligations, collocations, semantic preferences/associations. As corpus linguists have observed, words are usually embedded in contexts, and rarely occur in isolation. To find out what a word means, we have to examine its contexts. It is the contexts that determine its meaning (for the interpretation of meaning, see Chapter 3). Only corpora containing directly observable language data enable us to have a proper discussion of contextual meanings of words. So there is no point in discussing the meaning of a word in isolation solely on the basis of researchers' intuitions, or 'highly individual choice and description', in Josephs' words (1995: 284).

Analysis of the affective lexicon of a language can reveal the classifications and interpretative categories imposed upon its speakers. Most of the discussions of cultural specificity of emotions, however, are impressionistic and therefore not very reliable. To give a solid and sound documentation of cultural specificity of emotion expressions we have to look at directly observable language data. Goddard also agrees that it is time to embrace open data-gathering methods, including 'recourse to corpora of naturally occurring texts' (1997: 156). As far as I know, Teubert (2000) is the first linguistic study on emotion by using corpus-linguistic methodology. In this study, by focusing on the semantic field of *trauer*, *sorrow* in German and English, he addresses the deficiencies of the dictionary approach and the conceptual ontological approach to bi- and multilingual lexical semantics, and proposes corpus linguistics and parallel corpora as a new potential and promising approach to this issue.

In his another study on emotion *When did We Start Feeling Guilty* (Teubert, 2004b), he claims that *guilt* as a feeling has not entered into the Western discourse before roughly the second half of the 19th century, on the basis of historical evidence from dictionaries, psychological and cultural anthropological researches, early modern literature and a corpus. *Guilt* is, as we know, fuzzy and ambiguous. It means both ‘the state of having committed a wrong’ and ‘a feeling, a state of emotion’. In this study, 4114 occurrences (excluding the occurrences referring to the first sense) of *guilt* as a feeling are found in the Bank of English. Its collocation profile contains words like *inferiority, remorse, inadequacy, shame, intense, terrible, anger, failure, stress, self, deep, parents, strong, burden, release, painful* etc. Here are some citations for *guilt* as a feeling extracted from the corpus (p. 154-157):

- Despairing, filled with remorse, guilt and self-pity, Kevin could not forgive himself.*
- The profound disruption in basic trust, the common feelings of shame, guilt and inferiority*
- profound feelings of failure, guilt, self-loathing, and shame for having given in*
- most females who are unable to control their eating behaviour feel intense guilt*
- shame and guilt are stuff that makes us hide ourselves*
- passions can be aroused by liberation from these guilts*

We can see that, in the first four citations above, *guilt* is caused by being unable to forgive oneself or to control one’s eating behaviour, by ‘a disruption of basic trust’ (p. 155), for giving in when one should not have. The last two citations talk about the consequences of *guilt*. We can also see that if *guilt* is conjoined with *shame, remorse, self-loathing* and the like, the purpose seems to be ‘less a differentiation of these feelings than an intensification of the overall effect’ (p. 155). These relevant context

words of *guilt*, along with the concordance lines (or contexts) in which it occurs, would help capture the meaning of *guilt* as a feeling.

Following Teubert, Kondo (2004) makes a comparative study of *unhappy* on the basis of a small Japanese-English newspaper corpus. Although her study focuses on syntactic patterns rather than meaning and the data from that small corpus is not sufficient to allow a real statistical analysis, it is still a meaningful pilot empirical study to show, in some respects, the similarities and differences of *unhappiness* in English and Japanese.

Apart from what has been mentioned, some other corpus-based studies on emotions can be found. For example, Stefanowitsch (2004) analyses the metaphors of *happiness* and *joy* in English and German based on news texts available via the web archives of newspapers using the search engine Webcorp. However, this study is based on the cognitive framework and focuses on the differences in metaphorical mappings between these two languages. It only uses the corpora to find the instances of metaphor, talking about nothing about frequency or collocates. This is the typical way most researchers make use of corpora. On the other hand, while the advent of parallel or translational corpora has rapidly revived contrastive linguistics, very few corpus-based contrastive studies deal with meaning or meaning differentiation. Most of the studies are associated with translation or usage.

A corpus-based semantic analysis of emotion items can provide a more accurate description of their meanings and lead to a better understanding of them. Cross-linguistic studies on emotions can also benefit from the corpus-linguistic approach. By

looking closely at how emotion items are used in various languages, we can observe both similarities and differences in the use of these emotion items and their equivalents in other languages; we can identify what elements they share with similar expressions from other languages and what elements are specific to their own language. An empirical examination of emotion items in large corpora is likely to give more objective results and to further the discussion, which until now has all too often relied on rather vague introspection or the analysis of interviews with a number of participants that can hardly be considered representative. My work with corpora ensures that my investigation leads to tenable results.

2.3 Contrastive lexical semantics

2.3.1 Lexicalisation

One important question that comes up frequently in lexical semantics is that of the motivation for lexicalisation. The question is what makes certain sorts of notions or semantic configurations get lexicalised as single words rather than expressed by phrases? In specific terms, when would somebody say *kill* rather than *cause to die*? For Fillmore, ‘the act of lexicalising something is to present it as an established category of human thought’ (2003: 258). If a lexical item exists in a language, in Fillmore’s words, ‘it must exist as some part of a frame and must correspond to some part of a schema’ (p. 258). His example of the English word *vegetarian* illustrates this point successfully. *Vegetarian* does not merely refer to a person living on vegetarian food, such as fruits, vegetables, grains. It implies that there are many people who eat meat, and by contrast

there are people who deliberately do not eat meat. Otherwise, there would be no need for a particular word to represent such a concept and hence such a distinction.

Different languages organise lexical information in different ways. Words from different languages can differ in semantic ‘transparency’, ‘motivation’, and ‘analyticity’ (Fillmore, 2003: 274). Ullmann (1958: 680, cited in Fillmore, 2003: 274) describes this as the dosage of arbitrary vs motivated words in a language. Fillmore (2003) presents a detailed analysis of this feature. He claims that ‘one language might systematically have large repertoires of unanalysed words with highly specific meanings; another might have elaborate systems of semantically motivated word-formation principles; a third might depend more on syntactic processes for expressing complex notions. Languages can be seen to differ strikingly in the relative weighting that is given to one or another of these devices’ (p.270). He finds that Chinese month-names are highly motivated, since they have the form ‘1-month’, ‘2-month’, etc., while the English month-names are arbitrary with only slight etymological motivation in September, October, November, and December. He makes a study of English verbs in the domain of walking on the basis of a large English-Japanese dictionary and finds that the Japanese translation consisted of the word *aruku* (walk) plus a number of adverbs. *Stroll*, *saunter* and *meander* were all translated as *burabura aruku*, *plod* as *tobotobo aruku*; *totter* as *yochiyochi aruku*, *waddle* as *yoroyoro aruku*, and *lumber* as *doshindoshinto aruku*. His observation is that Japanese tend to combine mimetic adverbs with basic verbs to produce meanings of various ‘walking’ that are realised in English as single verbs, which share the the semantic core of ‘walking’ and differentiate from each other in

manner and attitude. Hence, Japanese and English vary systematically in their lexicalisation patterns, with Japanese more analytic in the domain of 'walking'.

Another related issue in contrastive lexical semantics is what Ullmann refers to as 'characteristic tendencies'. Take the expression of motion and location for example. Although all languages seem to decompose motion events into components such as manner, path, the moving figure, etc., they differ both in how they combine these notions into words (Talmy, 1985) and in how they categorise spatial relations they distinguish (Bowerman, 1989, 1996; Casad & Langacker, 1985; Lakoff, 1987). For example, Choi & Bowerman's study (1992) shows that English and Korean differ in the way by which they lexicalise the components of motion events. English uses the same verb conflation patterns for both spontaneous motion and caused motion, and path is encoded separately with the same path marker. The former is usually expressed by an intransitive verb and the latter a transitive verb. Korean, by contrast, uses different lexicalisation patterns for spontaneous motion and caused motion, and in most cases it uses different path markers for the two kinds of motions.

2.3.2 The problem of decomposition

It is essential to establish equivalence of lexical items across languages in lexical contrastive semantics. As mentioned earlier, there has been nearly a consensus that equivalent lexical items in different languages can be decomposed into identical set of features, which are referred to as conceptual primitives or semantic universals. The reduced concepts can thus serve as *tertium comparationis* for establishing cross-

linguistic lexical equivalence. However, taking this position poses problems. Krzeszowski (1990) presents a detailed analysis of this difficulty. To illustrate his point, he uses the example involving the words *man*, *boy*, *woman*, and *girl*, which are usually decomposed into the features [male], [female], [young] and [adult]:

	[male]	[female]
[adult]	man	woman
[young]	boy	girl

He puts forward four reasons for the failure of this sort of framework to serve as *tertium comparationis* and presents strong arguments against decomposition by giving a series of illustrative examples. Firstly, the features obtained from the decomposition analysis are not primitive at all. They can be further decomposed. He exemplifies this point by discussing the feature [young], which, in *The Concise Oxford Dictionary*, is defined as ‘that has lived a relatively short time; not mature or fully developed’; *mature* is defined as ‘complete in natural development or growth; fully developed in body and mind’, and *develop* is defined as ‘to grow into a fuller, higher, or maturer condition’. Here there is circularity, which is inherent in dictionary definitions. Features like [young], if they are expressed in natural languages, are problematic as *tertium comparationis*, since the words themselves are polysemous and ambiguous. Apart from being not mature, *young* has other senses, such as ‘newly begun or formed; not advanced: The evening is young’, ‘pertaining to or suggestive of youth or early life: young for her age’, ‘vigorous or fresh, youthful’ etc. The relatedness among its senses does not necessarily lead to the exact correspondence between the lexical items from different languages. For example, *young* is the equivalent to the Polish adjective *mlody* in its sense ‘that has lived a relatively

short time', but in the sense 'being in the early period of growth', it is equivalent to *swiezy* in Polish.

Secondly, the features resulting from the decompositional procedure fail to 'reveal accurately the meanings of the analysed lexical items insofar as they leave some area of the meaning unaccounted for' (Krzyszowski, 1990: p.87). The well-formed *young man* and *young boy* seem to demonstrate that *boy* does not only mean 'young male' and the description is inaccurate. Thirdly, in lexical contrastive studies, the combination of features listed in matrices is insufficient to provide a complete account of the meanings of the lexical items compared. The matching of features [young] and [man] will lead to such items as *boy*, *lad*, *youth*, *urchin*, and *master*. A host of other features, especially emotional aspects of meaning of such words, is needed to draw distinctions among them. Fourthly, the componential analysis cannot deal with various fuzzy concepts successfully. 'This procedure cannot be applied to various concepts which have no clear-cut boundaries and which shade into other concepts in a non-discrete way' (p.88). Krzyszowski uses a set of words expressing non-discrete concepts to illustrate this point: *forest*, *wood*, *woodland*, *grove*, *copse* and *spinney*. Krzyszowski says that the differences between *spinnery*, *grove* and *wood* in *The American Heritage Dictionary* are 'a matter of degree' (p.88), and the adjective *small* denotes a relative concept: *spinnery* is glossed as 'a small grove', *grove* as 'a small wood', while *wood* as 'a dense growth of trees; forest'. Such features cannot be defined in terms of the classical view of 'sufficient and necessary conditions' on which the componential

analysis is based. It is not a yes-no matter, which can be answered simply by finding out whether the defining elements of a certain category are present.

It seems that the establishment of lexical equivalence across languages, by decomposing words into primitive concepts, is more than difficult to achieve because the idea of primitive concept itself is problematic since we have to express it in a natural language, where the great majority of words are polysemous and ambiguous, and traditionally their definitions rely on other words in dictionaries, which will unavoidably lead to circularity. As can be seen, meaning is far more subtle than a component analysis based on discrete features can account for, and far more complex than any print dictionary can describe. Meaning is often fuzzy and difficult to define satisfactorily. James (1980: 168) suggests that, for a cross-linguistic lexical study, a better idea is to start with 'a perceived or assumed similarity' between contrasted items rather than to seek primitive or atomic units. We can only refine the the assumed similarity by analysing them thoroughly and by adding explicitness and precision, but we cannot expect to deconstruct them to atomic units and then reconstruct them like building blocks. The human mental world is not as clear and orderly as the physical one. Even if we can deal with satisfactorily lexical items that are related to physical behaviour, such as 'walking' verbs, by decomposing them, we cannot say it applies to all domains of language.

As a matter of fact, any aspect of lexical semantics can serve as a basis for cross-linguistic comparison. We thus obtain contrastive studies of various lexical sets such as colours, kinship, modality, or vision (see Krzeszowski, 1990 for details of these

studies). It is true that traditional contrastive studies have, to some extent, yielded some research findings. However, such investigations are incomplete and inadequate because they are not supported by directly observable quantitative data and their descriptions are largely based on individual researcher's introspection, dictionaries or scanty empirical evidence. Due to the inadequacies of the traditional approach, there has been a growing need to employ corpus methods for contrastive lexical semantic studies. This modern methodology arrives at the meaning of a word by focusing on the linguistic behaviour it exhibits in various contexts, namely, concordance lines and collocates, rather than by seeking its semantic primitives. Distinctions of semantically similar lexical items within the same language or across languages can be drawn without decomposing them. The idea is that to know a word does not necessarily mean that you need to decompose it; to understand a foreign word does not mean you need to give up your own language and look for language-independent semantic atoms. What you need to do is to figure out its position in its own language, and then compare it with your own.

2.4 Corpus linguistics, parallel corpora and cross-linguistic research

2.4.1 Corpus linguistics

Corpus-based analyses are 'empirical, utilise a large and principled collection of natural texts, make extensive use of computers for analysis using both automatic and interactive techniques, and depend on both quantitative and qualitative analytical techniques' (Biber, Conrad, & Reppen, 1998: 4). When we incorporate corpus linguistics, the only object of analysis will be delimited to directly observable language data. In Sinclair's (1991: 4) words, 'the ability to examine large text corpora in a systematic manner

allows access to a quality of evidence that has not been available before' (Sinclair, 1991). We can explicitly express the criteria for including texts in a corpus to make a representative sample of the discourse in question, thereby producing replicable corpus evidence. The corpus-based observations and results become transparent and verifiable regardless of the differences in interpretation. Corpus data is 'observed' rather than 'elicited'; it is superior to 'causal citational' data, which is also observed data, in that it is 'systematically and coherently organised' (Ooi, 1998: 51).

According to corpus linguistics, the meaning of a word cannot be viewed independently of the contexts in which it occurs. 'Lexical meaning exists solely in the universe of discourse as the complex web tying words to other words via the contexts in which they occur and tying words within their contexts to all other occurrences of these words in all preceding texts' (Teubert, 2000: 168). Looking at the context, we will find that it is often a specific co-occurrence pattern that disambiguates different usages or meanings of a word. Words can only be understood in their contexts and collocations. The traditional approach to bilingual lexical semantics is the bilingual dictionary, which consists of lexical items in one language and their equivalents in another, usually based on the linguistic intuitions of lexicographers. The information provided in the dictionary is often not explicit and lexical items are treated in isolation. That is why bilingual dictionaries are always insufficient for matching up lexical items across languages.

2.4.2. Translation equivalence, parallel corpora and cross-linguistic research

Translation equivalence is an essential issue in contrastive lexical studies. It is ‘not something that latently always exists and just has to be discovered. Translation equivalence has to be construed. As with meaning, this construal is a communal activity, only it does not involve a discourse community of a specific language such as English, but the community of bilingual speakers of the two languages involved. One translator will come up with a proposal, which is then negotiated with the other members of that community, until agreement is reached and every translator starts using the same equivalent, or until several equivalents are considered acceptable and translators choose among them’ (Teubert & Cermakova, 2004: 146-147). So translation equivalence as a discourse construct does not exist independently of languages use in discourse, i.e. translation process.

The terminology of ‘parallel’ corpora is a problem as the terms ‘parallel’ and ‘comparable’ corpora have been used in quite different ways, the former referring both to translation corpora and to comparable corpora, the latter to monolingual corpora of comparable texts as well as to corpora of comparable texts in different languages. In this study, following Teubert (1996), I use *parallel corpus* to refer to a corpus composed of source texts and their translations in one or more different languages, and *comparable corpus* to one which is composed of L1 data from different languages on the basis of same sampling techniques.

‘Parallel corpora are repositories of the practice of translators. The community of translators from language A to language B and vice versa know a lot more about

translation equivalence than can be found in any of the bilingual dictionaries for these languages' (Teubert & Cermakova, 2004: 155). Translators often do not rely on bilingual dictionaries and 'have to acquire a competence that is the result of experience and interaction with other members of the bilingual discourse community of which they are a part' (Teubert & Cermakova). Parallel corpora can not only tell us how languages differ but also can provide insights into the functioning of the languages involved. They can provide more precise information about the co-occurrence patterns and frequency data of the source items as well as of semantically related expressions which are brought into the picture by looking at their translations. The relative frequencies of semantically similar items in the contrasted language pair may relate to their semantic and pragmatic ranges of senses and functions. In other words, obtaining the quantitative information about semantics and translations is crucial in making sense of the results. The task is by no means simple, as the results at first sight often look 'messy'. But once the picture starts forming itself, we can see language in a novel way.

Parallel corpora are viewed as an important resource for contrastive studies. Their usefulness has been amply testified by empirical studies (Aijmer, 1998; Altenberg, 1999; Santos, 1998; Schmied, 1998; Xiao, McEnery, & Qian., 2006 etc.). For example, Santos' study shows that there is a great difference between English and Portuguese perception verbs in frequency, the way they are used, the kind of objects they co-occur with and the kind of information that is co-specified with them. Translation data can shed new light on the similarities and differences between the compared language pair. At the same time, the contrastive focus, by revealing what is general and what is

language-specific, can give a fuller picture of a given phenomenon in each of the language pair than a mono-linguistic study would do (Aarts, 1998; Santos, 1998; Viberg, 1998). In other words, translation data not only provide a great impetus to contrastive studies, but also offer a fresh perspective on mono-linguistic studies. Further, they ‘not only lend themselves to descriptive and applied approaches, but are also suitable for theory-oriented studies’ (Aarts, 1998: x). The range of linguistic research fields covered by the parallel corpus approach is very broad: semantics, lexis, syntax, discourse, translation process and translation style. Aijmer and Altenberg (Aijmer & Altenberg, 1996: 12) outlined some potential uses of multilingual corpora for cross-linguistic research:

- They give new insights into the languages compared—insights that are likely to be unnoticed in studies of monolingual corpora;
- They can be used for a range of comparative purposes and increase our knowledge of language-specific, typological and cultural differences, as well as of universal features;
- They illuminate differences between source texts and translations, and between native and non-native texts;
- They can be used for a number of practical applications, e.g. in lexicography, language teaching and translation.

However, the idea of using parallel corpora for cross-linguistic research is challenged by some linguists due either to the translator’s idiosyncrasies, or to ‘translationese’ (for a detailed discussion of ‘translationese’, see Gellerstam, 1986, 1996), namely the deviance in translated texts induced by the transfer of phenomena from the source to the target language (Johansson & Hofland, 1994) (for a detailed discussion of the problems in the use of translation corpora, see Johansson & Oksefjell, 1998). For the reason of translationese, Altenberg & Granger (2002: 17) provide the following:

Translators transfer texts from one language or culture to another and the translations therefore tends to deviate in various ways from the original. We have already mentioned possible translation effects---traces of the source language or universal translation strategies—and they involve additions, omissions, and various kinds of ‘free’ renderings that are either called for or motivated by cultural and communicative considerations.

It seems that translationese is inherent in translations (Baker, 1993; McEnery & Wilson, 2001; McEnery & Xiao, 2002; Teubert, 1996) since it is ‘motivated by cultural and communicative considerations’, but it does not follow that translated texts cannot be used for cross-linguistic research (Aston, 1999; Baker, 1993; Johansson & Hofland, 1994; Mauranen, 2002; Teubert, 1996 etc.).

Despite translationese, it is not totally impossible to establish equivalent lexical links across languages by using parallel corpora. Of course, the solution to this problem is not simple. Obviously, we cannot trust all the translations in the corpus as target equivalents. To determine which candidates can be regarded as ‘equivalents’, one choice is to resort to the procedure of ‘back-translation’ (Ivir, 1987), i.e. to look for those in the target language that can be translated back into their original forms in the source language. This procedure will filter out those unreliable equivalents arising from the translator’s style or other translational strategies, but, of course, it requires a bidirectional parallel corpus, i.e. a corpus containing original and translated texts for both directions. Another way is to look for the more recurrent translation equivalents on the basis of quantitative information (Krzyszowski, 1990). If the same translation pattern occurs recurrently in a parallel corpus, it obviously increases its possibility to be viewed as a translation equivalent. A corpus which contains texts from a variety of translators can be trusted. The third choice would be to combine the above two---back-

translation and frequency--- to arrive at the 'mutual correspondence' (also referred to as translatability) of two items in a bidirectional parallel corpus (Altenberg, 1999). Cross-linguistic equivalence, like any notion of equivalence, is a matter of judgment (Chesterman, 1998), so the notion of equivalence is only a relative concept. In practice, 80% will be a high value (see Altenberg, 1999). The value obtained often serves as a statistical measure of the degree of correspondence between the compared items (Altenberg).

Dyvik (1998) argues that translation corpora can enrich linguistic semantics not only because they incorporate a multi-lingual element into this kind of research, but also 'because the activity of translation is one of the very few cases where speakers evaluate meaning relations between expressions without doing so as part of some kind of meta-linguistic, philosophical or theoretical reflection, but as a normal kind of linguistic activity' (p.51). Teubert claims that 'the common practice of translators for a given language pair reveals the conventions of this bilingual language community. There is no reason to assume that the conventions of a monolingual language community concerning the meanings of text segments differ in principle from those of a community of translators' (1999: 23). This new understanding of translations greatly reduces our worries about translationese, especially when we deal with meaning.

Due to the well-known deficiencies in previous contrastive studies, there has been an increasing awareness that cross-linguistic studies cannot rely on introspection or dictionaries, but must be firmly based on naturally occurring language data. The development of machine-readable corpora and relevant computer technology are well-

suiting to meet these new demands. The corpus-linguistic approach to cross-linguistic research allows us to investigate with more objectivity and greater accuracy how words are used in natural or at least relatively natural contexts. Parallel corpora have been successfully used in a variety of cross-linguistic studies, though most of them are concerned with non-meaning aspects of language.

My research is going to focus specifically on lexical meaning, so large corpora are required to fulfil this purpose. To find what a word means, we will search it in a monolingual corpus and examine the words in its company. A large corpus will yield sufficient citations of the examined word. We can set up its collocation profile, which, coupled with paraphrases and examples, will tell us what it means. The parallel corpus will be used to link the synonymous words from the two languages. After arriving at the meanings of the words in monolingual corpora respectively, we can compare and contrast the semantically equivalent word(s) linked by the parallel corpus.

2.5 Summary

The practice of taking English emotion terms like *sadness*, *unhappiness*, *anger* and *fear* as labels for universal psychological states has been argued to be ethnocentric and problematic (Wierzbicka, 1986, 1992b, 1999). Even the traditional psychological categories, such as 'emotion', 'sensation' and 'cognition', have been claimed to be only cultural artefacts of western societies (Harre, 1986; Lutz, 1988; Wierzbicka, 1999). So far, psychology has been working within the circle of shallow ethnocentrism. Despite the obvious weaknesses mentioned earlier, psychological matching experiments are

often taken as proof that all languages have words for the basic emotions. Other techniques used by anthropologists and cross-cultural psychologists include similarity judgment tasks, sorting tasks, listing tasks and rating tasks. All of these experiments face the problem of translation since they all depend on elicitation and/or interpretation in natural language. In my view, it is time for psychology to embrace new approaches to deal with the problem of semantic variation, rather than simply ignore it, because the new approaches will lead to fresh research findings—the so-called ‘basic’ emotions are not basic at all and human emotions are not so universal.

In the field of linguistic semantics, researchers are usually well aware of the problem of ethnocentrism and the difficulties of translation. The NSM framework is trying to demonstrate the cultural-specificity of emotions through so-called cultural-universal concepts. However, it is problematic whether the notion of ‘semantic primitives’ itself can hold water since they are English words themselves. As noted earlier, it is questionable that Wierzbicka’s universal concepts like ‘I’, ‘good’ and ‘bad’ are truly universal across all languages. Furthermore, it is much too simplistic to think the description of the world can be broken down to 60 basic universal concepts. It is hard to imagine words with complicated meanings can be defined or explained by these so-called semantic primitives. Besides, Wierzbicka is only concerned with words in isolation and focuses on their ontological meanings without looking at its contexts, collocations, colligations, semantic preferences/associations. The analyses within the NSM framework are still mainly based on researchers’ introspection rather than directly observable language data.

The corpus-based approach furnishes cross-linguistic lexical semantics with a valuable new analytical and descriptive perspective by basing its analyses on naturally occurring language and regarding word meaning as the sum of its typical contexts, including its collocations, colligations, semantic preferences/associations and paraphrases. This approach examines a word as a whole without trying to discompose it. So it does not have the language neutrality problem the NSM theory faces. It offers not only more evidence, but also a 'new approach'.

CHAPTER 3

THEORETICAL FRAMEWORK

3.1 Corpus linguistics: Theory or methodology?

It has been a great controversy whether corpus linguistics is only a methodology that can be used in almost any area of linguistic research, or whether it is an independent discipline in its own right in the same sense as cognitive linguistics. It is well acknowledged that corpus linguistics has opened up exciting possibilities which could not have been envisaged without corpora; in other words, the corpus technology has led to new findings which cannot be achieved by intuition alone. But the real strength of corpus linguistics lies in the fact that these findings have shaken the foundations of linguistics (Scott & Tribble, 2006). For example, ‘Sinclair’s (1991) corpus-based questioning of the parts-of-speech which were so traditionally established, or Francis, Hunston and Manning’s (1996) pattern grammar arguing that grammar is much more local than it had seemed’ (Scott & Tribble, 2006: 4). The wide recognition of pattern grammar suggests that lexis has shaken the dominant position of syntax and grammar in language study and attracted considerable attention (Scott & Tribble). Tognini-Bonelli (2001) proposes that a distinction should be drawn between the corpus-based and corpus-driven approaches. In the former, corpora are used only to ‘expound, test or exemplify theories and descriptions that were formulated before large corpora became available to inform language study’ (p. 65), while in the latter, the aim is to ‘derive linguistic categories systematically from the recurrent patterns and the frequency

distributions that emerge from language in context' (p.87). 'The distinction between 'corpus-based' and 'corpus-driven' approaches may be seen to correspond to the 'top-down' and 'bottom-up' approaches' (Ooi, 1998: 52). According to Tognini-Bonelli (2001), only the corpus-driven approach is related to corpus linguistics as a separate discipline.

'When corpus linguistics started off, the main reaction was the excitement about various unexpected findings and technical possibilities. However, the interest in the theoretical consequences of corpus linguistics has been gaining importance' (Mahlberg, 2005: 17). Corpus studies have yielded numerous findings on the following topics existing theories have either failed to explain or totally ignored: (extended) unit of meaning or lexical item (Sinclair, 1996, 1998, 2004; Sinclair & Teubert, 2004; Stubbs, 2009; Teubert, 2005 etc.), pattern grammar (Francis, et al., 1996; Hunston & Francis, 1998, 1999), collocation, semantic prosody (Bednarek, 2008; Hunston, 2007; Louw, 1993, 2000; O'Halloran, 2007; Ooi, 1998, 2000, 2008; Partington, 2004; Sinclair, 1996, 1998; Stewart, 2010; Whitsitt, 2005; Xiao & McEnery, 2006 etc.), part-of-speech (Sinclair, 1991), lexis as the core of language (Hoey, 2004, 2005; Sinclair, 1991, 2004; Teubert, 2004a, 2005 etc.), the relationship between form and meaning (Hoey, 2005; O'Halloran, 2007; Sinclair, 1985, 1987, 1991; Stubbs, 1996; Teubert, 2005; Teubert & Cermakova, 2004; Tognini-Bonelli, 2001; Williams, 1998 etc.), paraphrase as interpretation of meaning (Teubert, 2005, 2010; Teubert & Cermakova, 2004). Only corpus-based analyses can discover them. 'The potential of corpus linguistics and its role in linguistics will be seriously underestimated if we regard it as a method only'

(Mahlberg, 2005: 17). Corpus linguistics is a promising enterprise as it has provided a new philosophical approach to linguistic enquiry (Tognini-Bonelli, 2001) and a new look at language (Teubert, 2005, 2010), which greatly help reveal the real nature of language.

Both 'corpus-based' and 'corpus-driven' approaches are useful notions. In this study, the corpus-based approach is taken to mean that the notion of 'lexical word' still applies and is not reformulated. Complementing this perspective, the 'corpus-driven' approach uses empirical evidence for reformulating the traditional category of 'lexical word' (which is not taken for granted). Words are polysemous, and so are emotion words. I will focus only on the 'feeling' sense for emotion words. In dealing with the data, this study is more corpus-driven as it is based on the corpus-linguistic framework and takes the data as the centre of the analysis. Despite the corpus-based vs corpus-driven distinction noted earlier, in practice a mixed approach of both is often necessary (Ooi, 1998). A more recent notion is 'triangulation' which refers to 'the use of more than one approach to a research problem, including using several research methods or instruments to investigate the same thing, so as to corroborate findings and minimise the possibility of a method distorting the evidence' (Sealey, 2010: 64). It seems that the combination of approaches to a research problem has become the current trend. We may use corpus data to triangulate our intuitions in order to make a more accurate description of the lexicon.

3.2 The corpus-linguistic view on meaning

3.2.1 Meaning as a social construct

The claim that meaning is a social construct is either articulated or implied in social constructivism (Berger & Luckmann, 1966; Lave & Wenger, 1991; Vygotsky, 1978), focusing on the construction in the individual mind derived from his/her interaction with social structures, and social constructionism (Burr, 1995; Gergen, 1994, 1999), emphasizing more on the collaborative construction through social interaction. Corpus linguistics is closer to social constructionism in terms of the collaborative construction of meaning (details will be addressed in the next section). Critical discourse analysis (CDA, henceforth) also works along this line. CDA, by describing, interpreting and explaining the situation, recipients, consequences and social responses of the discourse, aims to reveal the underlying ideological bias of the discourse and uncover inequality and social reality (Fairclough, 1989). CDA and corpus linguistics differ in their attitude towards the relationship between society and discourse (Teubert, 2010). CDA (Fairclough, 1989) claims that it is society that determines discourse, while corpus linguistics (Teubert, 2010: 121) believes that 'it is the discourse that constructs social structures, not the other way around'. Despite the theoretical difference stated above and CDA's heavy focus on the relation between ideology and inequality in social reality, CDA confirms the claim that meaning is constructed in discourse by language users. It has done much in revealing how meaning/ideology is constructed to maintain the unequal power relations in society. However, there is no standard methodology for critical discourse analysts to follow. Stubbs (1997) raised some practical questions such

as how frequency of use relates to naturalization, how the data is selected and if they are representative. Corpus linguistics opens up a new battlefield to deal with meaning construction in discourse. The corpus-linguistic approach can circumvent CDA's disadvantages noted above because corpus-based analyses are empirical and 'depend on both quantitative and qualitative analytical techniques' (Biber, et al., 1998: 4).

The corpus-linguistic claims relevant to the negotiation or construction of meaning from Teubert's 25 theoretical theses are listed below (2005: 2-4):

1. The focus of corpus linguistics is on meaning. Meaning is what is being verbally communicated between the members of a discourse community. Corpus linguistics looks at language from a social perspective. It claims no privileged knowledge of the workings of the mind or of an innate language faculty.
3. Every text segment, word, multi-word unit, phrase etc., can be viewed under the aspect of form and the aspect of meaning. The form is what represents the meaning, and there is no meaning without the form by which it is represented. Text segments are symbolic; they always mean something to someone. Normally the members of the language community deal with text segments without being aware what these text segments mean just as people are often jealous without being conscious of their jealousy. Unless there is some (potential) communication disorder, there is no need to discuss meaning within the discourse community.
4. Meaning is in the discourse. Once we ask what a text segment means, we will find the answer only in the discourse, in past text segments which help to interpret this segment, or in new contributions which respond to our question. Meaning does not concern the world outside the discourse. There is no direct link between the discourse and the 'real world'. It is up to each individual to connect the text segment to their first-person experiences, i.e. to some discourse-external ideation or to the 'real world'. How such a connection works is outside the realm of the corpus linguist.
5. For corpus linguistics, the meaning of a text or of a text segment is independent of the intentions of its speaker (its author). The dislocation of the speaker/author from his or her text distinguishes written (recorded) language from spoken language. In spoken language, the speaker is usually present, and if there is a communication disorder, we ask: 'What do you mean?' and not: 'What does this mean?'

7. Corpus linguistics makes general and specific claims about the discourse, based on the analysis of a suitably selected cross-section of it, i.e. the corpus. General claims have to do with rules or with probabilistic expectations. They fall within the field of grammar or variation or language change, and also into the field of lexical meaning insofar as a text segment occurring in a text can be viewed as an instantiation (a token) of a lexical item. Specific claims are interpretations of texts or text segments viewed as unique occurrences.

From the above theses, we can see that, according to Teubert (2005), corpus linguistics focuses on meaning in the collective sense as found in a discourse, not on understanding in the individual sense or intention as found in people's heads. Through the explication of a haiku randomly produced by a computer program, Teubert (2010: 215-240) shows that the meaning of a text is separate from the intention of its author. The general claim in Thesis 7 refers to the meaning in the collective sense, while the specific claim is the interpretation based on the collective sense, regardless of what is intended by its author or speaker. Meaning emerges as social constructs independently of the discourse-external reality; not only meaning, but our perception of reality is also constructed through social interaction and mediated by the use of language. This perspective is very much in line with constructionism (Gergen, 1994, 1999).

Numerous corpus studies show that each linguistic form has a certain meaning (Hoey, 2005; O'Halloran, 2007; Sinclair, 1985, 1987, 1991; Stubbs, 1996; Tognini-Bonelli, 2001; Williams, 1998 etc.). This seems to suggest that the members of a discourse community tend to attach a certain meaning to a certain linguistic form. Linguistic forms are only linguistic symbols to represent meanings. Form and meaning cannot be separated. The integration of meaning and form is the outcome of social construction. Members in a discourse community take part in the negotiation or

clarification of the meaning of a given linguistic unit in progressive discourse. Words or phrases are just social constructs to which discourse members attach meanings in social contexts. They have nothing to do with the discourse-external reality. In a recent interview (R. Zhang, 2009), Teubert exemplifies this view:

One of the basic units of biology is the species. But today all biologists would admit that what we call a species is not something that exists out there in the real world; what we call a species is the result of never-ending negotiations within the discourse of expert biologists. A species is a discourse concept, nothing more. We can find it useful today and discard it tomorrow, just as the infamous phlogiston has been discarded after it had outlived its usefulness. The idea that science can tell us the truth about natural phenomena is not helpful. There is no observation without knowing first what to observe, and this 'knowledge' does not refer to 'true' properties of things in the 'real' world out there, they are nothing but agreements accepted by certain discourse communities (p. 87).

According to Teubert (2005, 2010), meaning is only a social construct and the reality is constructed via discourse; the discourse-external reality is meaningless because it cannot be discussed, exchanged, shared and known; the personal experience will enter the discourse after it is shared and discussed. We come to know the world only through all the texts (written or spoken) exposed to each of us, which vary greatly from person to person. It is impossible to get access to the individual discourse, i.e. all the texts one individual has been exposed to, so we can only try to reveal the linguistic truth based on the public discourse in the form of corpora, which is also incomplete because the great majority of the texts produced by us have been lost. The meaning of a linguistic unit is neither what it refers to out there in some discourse-external reality nor some representation people have in their minds; meaning is usage and paraphrase (Teubert & Cermakova, 2004) as we find in discourse.

3.2.2 Meaning as paraphrase

Meaning can be rephrased in different words. Corpus linguistics takes meaning as paraphrases in discourse. Teubert's theses relevant to this claim are shown below (2005: 6-7):

19. Meaning is paraphrase. Whenever lexical item tokens are the cause of a communication disorder, their meaning will be negotiated, described or explained, replaced by synonyms, and sometimes even 'defined' as in dictionaries or in encyclopaedias. What we find paraphrased in the discourse are text segments that are looked upon intuitively by the members of the discourse community as units of meaning. However, the same lexical item type can be paraphrased in an infinite variety of ways. Therefore, whenever a lexical item token is being paraphrased, we can view it from two perspectives: as an instantiation of the lexical item type, and as a unique occurrence. From a synchronic perspective, the meaning of a lexical item type is a generalisation on all the paraphrases we find for the instantiations of this lexical item. But paraphrases are also relevant for specific claims. From a diachronic perspective, it is the history of paraphrases of a recurrent text segment, as evidenced in its intertextual links, that tells us what it means as a unique occurrence.
20. There is no true and no fixed meaning. Everyone can paraphrase a unit of meaning however they like, therefore the meaning of any lexical item type is always provisional. The next paraphrase may already lead to a revision. The members of the discourse community will continue to negotiate, among themselves, what a unit of meaning means. They may agree or not: the issue is not truth, but acceptance. An explanation, a paraphrase that is widely accepted and re-used, is more relevant than a paraphrase that is never repeated, just as texts which are constantly referred to are more relevant than texts that leave no traces in subsequent texts.
21. The discourse is a self-referential system. Natural language is the only codification system in which the functions of its elements are determined not by ascription from outside but by discourse-internal negotiation. This sets natural languages apart from formal calculi, like the code of mathematics.

Before discussing the above quotes, it is helpful to clarify some corpus-linguistic terms.

In corpus linguistics, 'lexical items' or 'units of meaning' can be 'single words, compounds, multiword units, phrases, and even idioms' (Teubert, 2005: 5). 'They are,

in principle, monosemous. This is what distinguishes the concept of a lexical item from the concept of a word' (Teubert). The above theses show that within a discourse individuals each develop their own interpretation of a given lexical item, in which some of them may be quite apart from each other, and this process is ongoing. Whatever is said about a given lexical item within the discourse contributes to the meaning of this lexical item, which is the accumulation of its previous occurrences, including what has been given in dictionaries. It is, from a synchronic perspective, a generalization of all the instantiations of this lexical item or a 'loose consensus', in Sinclair's words (Sinclair & Teubert, 2004), we are searching for or can achieve. The meaning of a given lexical item is co-constructed by the members of the discourse community, regardless of whether it is true or not. 'Lexical items and what they stand for are discourse objects (and not objects of the "real world") constructed through the contributions of the members of the discourse community' (Teubert, 2005: 7). Once we accept this view, the traditional distinction between lexical meaning and encyclopaedic meaning vanishes.

Paraphrases are interpretations in the form of explanations, explications, or redefinitions. An interpretation can be re-discussed, reinterpreted or further elaborated in the discourse. On this account, the meaning of a lexical item is co-referential within the discourse. Whenever a new lexical item is introduced into the discourse, it will be explained or interpreted in various ways. Its meaning will be argued and negotiated by its discourse members, when it is referred to numerous times later. It is through this negotiation and argument that its meaning is constructed socially and collaboratively within the discourse. A lexical item is a type and each instance of it is a token. The

contribution of each instance to its meaning varies considerably. Some instances are real paraphrases, like ‘sadness is a less strong feeling’, which can make obvious and significant contributions to its meaning, while others may only contain a piece of knowledge concerning the given lexical item that can tell us something about its meaning, and still others may make little contribution. Typical examples can contribute quite a lot to its meaning, so they can be taken as implicit paraphrases. Real paraphrases can be identified and also extracted by using corpus tools. Cheung (2009), based on the analysis of syntactically defined patterns in which paraphrases occur in texts, developed a method to automatically retrieve paraphrases from a corpus. However, sometimes, sufficient real paraphrases can not be identified in the corpus examined due to its limited size, which is seemingly quite big but actually still pretty small compared with the size of the human output. In such cases, we may need the help of implicit paraphrases, i.e. typical examples, to discuss meaning.

3.2.3 Meaning as usage

According to Teubert & Cermakova (2004), apart from paraphrase, the meaning of a lexical item is also its usage as found in the discourse, such as its frequency, contextual information, collocational profile etc., which can be established by a computer. Usage is ‘something we have to cope with as members of the discourse community’ (p. 128). This view is also shared by Hoey who argues that ‘the meanings of a word will have to be interpreted as the outcome of its primings, not the object of the primings’ (p. 81). In other words, the meaning of a word (sense) lies in the colligations, collocations and

semantic associations a word is primed for, not elsewhere. Hoey's notion of lexical priming will be addressed in detail at a later point in this chapter.

The present work is to explore how meaning can be described, defined, explained, negotiated and elaborated by the members of a discourse community. Based on the principles stated above, I will see how the meaning of a lexical item can be arrived at by examining its semantic profile based on all its instances in a corpus and also by investigating its real paraphrases and typical examples. Available real paraphrases differ greatly. For some lexical items, it is pretty difficult to find sufficient real phrases for the reason mentioned in Section 3.2.2, especially for Chinese ones. So in many cases, typical examples may work as implicit paraphrases to help disclose the meaning of the lexical item under question. In the analysis, I will not distinguish them clearly and refer to them all as examples. Another thing that should be noted is that I will use the principles discussed to guide my research at the macro-level, but I may not use these terms, such as social construct, paraphrase or usage, in my analysis. In presenting the data, I will use the corpus-linguistic structural categories of the lexical models at the micro level to direct my analysis. While most of the items I am going to examine in this research are lexical words, I will only focus on its sense as a feeling, so it is monosemous in this sense. That is why I identify them with lexical items.

3.3 A lexical model: colligation+collocation+semantic association/preference

3.3.1 Sinclair's notion of (extended) unit of meaning & his lexical model

In Firthian and neo-Firthian linguistics, the notion of word as the unit of meaning has been abandoned by most corpus linguists since a common word of a language typically

can have several meanings (Sinclair, 1991, 1996, 1998, 2004; Sinclair & Teubert, 2004; Stubbs, 1996; Teubert, 2004a, 2005; Tognini-Bonelli, 2001). Sinclair points out that a unit of meaning is 'rather a phrasal unit than a word' (Sinclair & Teubert, 2004: xx) as the co-occurrence tendencies between words are so strong that 'we must widen our horizons and expect the units of meaning to be much more extensive and varied than is seen in a single word' (Sinclair, 2004: 39). He argues that the model of a lexical item, or a unit of meaning, is the only one to facilitate the study of meaning. His classical examples *the naked eye*, *true feelings* illustrate what could be a unit of meaning. Sinclair focuses more on co-selection between words when he examines this concept. He explores extended units of meaning by trying to incorporate all the constraints around, for example, *the naked eye*, both at the lexico-grammatical and the semantic levels, into the extended unit of meaning, including the semantic preference of 'visibility', realised by strong collocations with words such as *see*, *visible* etc. He also argues that a lexical item is prone to variation; the canonical form would be the prototype of a lexical item. For example, *get in touch with* is the canonical form of the lexical item. There are all sorts of other verbs that could be substituted for *get*: *bring*, *be*, *keep*, *remain* etc. (2004). For Sinclair, lexical items or units of meaning have a strong tendency to become delexicalised. With regard to this, he states:

The meaning of words chosen together is different from their independent meanings, they are at least partly delexicalised. This is the necessary correlate of co-selection. If you know that selections are not independent, and that one selection depends on another, then there must be a result and effect on the meaning which in each individual choice is a delexicalisation of one kind or another (2004: 20).

This accounts for why a lexical item means different from the phrase which is interpreted according to the *open choice principle*.

Sinclair (1996, 1998, 2004) proposes a lexical model in which four structural categories are combined to describe lexical items---collocation, colligation, semantic preference and semantic prosody. In this model, the category of semantic preference refers to the similar semantic features shared by the item's collocates. Colligation is 'one step more abstract than collocation' (Sinclair, 2004: 32) and it has to do with the co-occurrence of grammatical choices. Collocation is defined as 'the occurrence of two or more words within a short space of each other in a text' early in his book *Corpus, Concordance, Collocation* (1991:170). In his famous article *The Search for Units of Meaning* (1996, 2004), Sinclair investigates *true feelings*, a common collocation that is not idiomatic in the normal sense, and summarises the result as follows (2004: 35):

- a semantic prosody of reluctance/inability
- a semantic preference of expression (and a strong colligation of a verb with the semantic preference)
- a colligating possessive adjective
- the core

Sinclair tells us that we choose this particular collocation when we talk about our reluctance to express our genuine emotions. This expression is almost never used in any other way. Many studies have sought to scrutinise the notion of semantic prosody (Louw, 1993; Partington, 2004; Whitsitt, 2005 etc.). The usual interpretation of semantic prosody is seen as too restrictive in that it is often associated with only three possibilities: positive, negative and neutral connotation. Sinclair uses this term in a less restrictive sense, i.e., it does not only refer to the above three possibilities, but refer to

any attitudinal/pragmatic meaning implied in the lexical item examined. Hunston (2007: 257), taking *budge* for example (Sinclair, 2004: 142-147), explicates Sinclair's idea on semantic prosody:

Sinclair's point is that the sequence '*inability+negative+budge* +(something)' is chosen in contexts where something difficult and important is being attempted, to no avail, and that the sequence '*unwillingness+negative+budge*' is chosen to express disapproval of someone's lack of flexibility. The sense of frustration is a more complex concept than a simple positive or negative evaluation, and it is clear that it belongs to the sequence as a whole, the unit of meaning, rather than just to the word *budge*.

From the above elaboration, it can be seen that semantic prosody, in Sinclair's sense, is used to define the discourse function of the whole (extended) unit of meaning. This explains why semantic prosody is usually not accessible via introspection alone, and is derived more accurately through a closer inspection of all instances of the lexical item in the concordance. This model is mainly intended for (extended) unit of meaning. Sinclair argues that the major structural categories involved in the model—collocation, colligation, semantic preference and semantic prosody---assume a central rather than a peripheral role in language description. He also argues that 'this model does not exclude single words that are apparently chosen on open-choice principles' (2004: 39). He analyses *place* (sense 13) to support his claim. However, this model 'has largely been illustrated by individual case studies, and it remains unclear whether all phrases, or only some, have semantic prosodies' (Stubbs, 2009: 12). If there is still difficulty in applying this model to larger (extended) units of meaning, it will certainly pose problems for single words, as they have a wider range of collocates which are less likely to be grouped under one single label for communicative purpose. To seek an appropriate

lexical model for my analysis of single words, let us look at Hoey's notion of lexical priming.

3.3.2 Hoey's notions of semantic association and lexical priming

In line with Sinclair's claim that grammar exists just because we can not say everything at once (Sinclair & Teubert, 2004), Hoey proposes a theory of lexicon which 'reverses the role of lexis and grammar, arguing that lexis is complexly and systematically structured and that grammar is an outcome of this lexical structure' (2005: 1). It is a theory of lexis moving from the surface of language product inwards into the psycholinguistic reality. Hoey is trying to illustrate lexicon through collocations, colligations and semantic associations 'that are built up in the individual mind, and harmonised in a collective consciousness of language' (Williams, 2006: 327) . His theory puts the lexis at the centre of language capacity and claims that grammatical and lexical choices will be governed by lexical primings. Hoey points out that the problem with the traditional theory is that it 'account[s] only for what is possible in a language and not for what is natural' (p. 2). Corpus linguistics 'is concerned with how naturalness is achieved and how an explanation of what is natural might impinge on explanations of what is possible' (Hoey). 'A key factor in naturalness, much discussed in recent years, is collocation' (Hoey). He also takes collocation as central to his approach; unlike Sinclair, however, he defines collocation as a psychological association between words.

Hoey's theory 'sets out to outline a new theory of language based on co-occurrence patterns that are stored in the individual's mind' (2005: 15). Some of his

claims are not entirely new, but they are expressed from a new psycho-linguistic perspective. Chomsky (1986) distinguishes between E-Language (externalised language) and I-Language (internalised language). Cognitive linguistics focuses on the latter, while corpus linguistics examines the former. Hoey's theory of lexical priming is intended to bridge the gap between the two (2005), with a focus on lexis rather than grammar. 'Priming can be seen as reversing the traditional relationship between grammar as systematic and lexis as loosely organised, amounting to an argument for lexis as systematic and grammar as more loosely organised' (Hoey, 2005: 9). Hoey puts forward a list of psycho-linguistic hypotheses using a corpus as evidence, of which those relevant to my research are shown below (2005: 13):

3. Every word is primed to occur with particular semantic sets; these are its semantic associations.
5. Co-hyponyms and synonyms differ with respect to their collocations, semantic associations and colligations.
6. When a word is polysemous, the collocations, semantic associations and colligations of one sense of the word differ from those of its other senses.

We can see that whether from a sociolinguistic, psycho-linguistic or purely linguistic perspective, whether taking a corpus as a discourse, a collective mind or a collection of texts, the corpus data tells us the same thing. One of the main differences between Hoey's lexical priming theory and Sinclair's lexical model lies in the notion of semantic association which replaces Sinclair's semantic preference and semantic prosody. Semantic preference refers to the habitual collocation of lexical items with linguistic expressions that come from certain semantic fields. Semantic associations allow collocates to be grouped into semantic sets. 'These reveal patterns of language that go

beyond collocation to form phraseological units with a string of variable elements' (Williams, 2006: 328). A simple comparison will reveal that the notion of semantic association is actually a different label of Sinclair's concept of semantic preference (cf. Hoey, 2004: 388; Ooi, 2008: 312). So the lexical model I will use for my research can be taken as an adapted version of Sinclair's model, with semantic prosody removed as it does not apply to every single word, or as Hoey's model, but from a sociolinguistic perspective, rather than a psycholinguistic one, as what I am talking about is the social construction of meaning. In analysing the corpus data, I will not distinguish them strictly as they both reveal the same thing.

Semantic preference is commonly shared by speakers of a given speech community (Partington 2004; Hoey, 2005). However, it varies in degree or strength (Louw, 1993; Partington, 2004); it is genre or register specific (Hoey, 2003, 2005; O'Halloran, 2007; Partington, 2004). In addition, different word forms of the same lemma may have distinct semantic preferences. For example, O'Halloran (2007) points out that in newspaper texts the verb *erupt* (in the word forms *erupted/erupts*) has a semantic preference for human phenomena whereas the noun *eruption* (in the word forms *eruption/eruptions*) has a preference for geological phenomena. In my research, I will examine various forms of a derivative family separately to reveal their respective differences in semantic preference.

Hoey's theory, to some extent, bears a similarity to Mel'c'uk's Meaning Text Theory and Fillmore's Frame Semantics. Williams (2006) makes the following remarks concerning this:

The discussion of semantic associations shows an analysis which comes close to the lexical functions of Mel'cuk and the Frame semantics developed by Fillmore, but both theories are overlooked. Meaning Text theory (Mel'cuk, 1988) and FrameNet (Fillmore, Johnson, & Petruck, 2003)² have a cognitive base and do attempt to build on the psychological underpinning of language. Whilst it is true that these theories have not grown out of corpus studies and are largely based on linguistic intuition, they both make use of corpora. These are largely corpus-based, to use the terminology of Tognini-Bonelli (2002³), that is to say the corpus is used as a test-bed to verify intuitions, as opposed to Hoey's corpus-driven methodology where the theory grows out of the exploration of the corpus.

In effect, lexical studies⁴ using corpus data as evidence to support their theoretical claims are not few (Chang, Chen, & Huang, 2000; Chief, Huang, Chen, Tsai, & Chang, 2000; Liu, 2002 etc.; Liu, Huang, Lee, & Lee, 2000; Tsai, Huang, Chen, & Ahrens, 1998; J. Wang, 2009 etc.). What is common among them is that various non-corpus-linguistic theoretical frameworks, mostly cognitive ones, are applied and corpora are merely used to provide data from which observational tendencies or generalization can be derived. But collocation, a central concept to corpus linguistics, is hardly mentioned in them. In this sense, it is better to refer to my research as corpus-driven because I will base my exploration on a corpus-linguistic framework, where collocations, colligations, semantic associations/semantic preferences and paraphrases will be put at the centre of the investigation. My findings will grow out of the corpora I use, not elsewhere. My aim is to gain a full understanding of the synonym pairs in the domain of sadness, tease out the subtle nuances of their meanings and differentiate them from each other.

² It is 'Fillmore et al, 2003' in the original text.

³ The original text makes a mistake in the year of the book, which should be 2001.

⁴ Corpora have been used in many lexical studies, but in few contrastive lexical semantic studies.

3.4 Summary

The present chapter has described the theoretical basis underlying this research, i.e. corpus-linguistic theories on meaning (as a social construct, usage and paraphrase) and a corpus-linguistic lexical model. This model originates from Sinclair's lexical model for (extended) unit of meaning. Hoey replaces semantic preference and semantic prosody with semantic association, a psycholinguistic notion, to adapt it for single words. The corpus analysis throughout this project is embedded in a corpus theoretical framework. The idea of the research is that the semantic information of a lexical item can be gleaned from the words in its company and its meaning can be teased out of its collocates, paraphrases and examples. If what is obtained from its collocates and paraphrases is not sufficient to reveal its meaning, which usually happens when its occurrence frequency in the corpus is too low, we have to resort to its examples. This research aims to confirm with corpus evidence that, from a corpus-linguistic perspective, 1) meaning is a social construct, the outcome of collaborative construction by the members of a discourse community; 2) understanding in people's heads based on intuition is private and individual and therefore has to be attested by empirical data, and meaning in the collective sense can only be found in discourse; and 3) meaning can be described in terms of usage, namely colligations, collocations and semantic preferences/associations, and paraphrases, both explicit and implicit.

CHAPTER 4

METHODOLOGY

4.1 Introduction

The present chapter deals with the methodological details for this research. Essential to corpus studies is the selection of the corpora on which generalizations are based and the corpora to be used in the present study are introduced in Section 4.2. Section 4.3 gives some brief information on the dictionaries and corpus tools. The methods of data processing and analysis are outlined in Section 4.4, in which some relevant issues such as comparability, representativeness, annotation are also addressed, along with some limitations of the methodology for the present work. In addition, Section 4.5.2.2.2 discusses at great length the existing problems in Chinese POS classification and tagging, which are closely related to the method adopted in this research.

4.2 Corpora used

4.2.1 The Bank of English

A comprehensive account of the latest version of the Bank of English (BOE, henceforth) is available in the *Release Notes 2009* at http://wordbanks.harpercollins.co.uk/Docs/WBO/WordBanksOnline_English.html.

Details will not be repeated here. What follows is only a brief introduction to some issues relevant to the current research. The BOE is a collection of samples of modern English language since 1960 and the bulk of the corpus ranges from 2001 - 2005. It contains a wide range of texts, both written and spoken, from various sources, including Britain, America, Australia, Canada, India, New Zealand, South Africa and Ireland.

Written texts mainly include newspapers, magazines, books, brochures, leaflets, reports and letters whereas the spoken part comprises transcriptions of everyday casual conversation, radio broadcasts, meetings, interviews and discussions. It covers a wide range of domains, such as news, tv-radio, fiction, lifescience, culture, biology, business, religion, computer, medicine, natural science, lifestyle, leisure and music etc. The vast majority of texts come from newspapers (51.8%). To be precise, it currently contains 455,039,614 tokens of word forms. The BOE has benefited a lot of corpus research on language, for example, Sinclair (1991), Renouf & Sinclair (1991), Hunston & Francis (1998, 1999), Moon (1998). It has been extensively used for the analysis of vocabulary, grammar and usage of English. This thesis will show in detail how the BOE could be used wisely to assist with a contrastive lexical analysis.

4.2.2 The CCL Corpus

The Chinese corpus compiled by the Centre for Chinese Linguistics (CCL), Peking University, China, consists of two sub-corpora: Old Chinese and Modern Chinese. I will only use the modern sub-corpus for my research, but for the sake of convenience, henceforth I will refer to it as the CCL corpus. It contains 307,317,060 Chinese characters⁵ (sourced from http://ccl.pku.edu.cn:8080/ccl_corpus/CCL_CC_Sta_Xiandai.pdf, updated on July 20th, 2009). It is a collection of samples of Modern Chinese (since 1919, which marks the beginning of Modern Chinese), both written and spoken. The great majority of the corpus ranges from 1993-2009, according to its documentation

⁵ According to its documentation available online, the CCL corpus contains 307,317,060 tokens and 9711 types. We may infer that it only counts Chinese characters as punctuation marks are not included in its list of character types.

(sourced from http://ccl.pku.edu.cn:8080/ccl_corpus/CCL_Xiandai.pdf). It contains a broad range of texts from Mainland China, Taiwan and Hong Kong, covering a wide range of text types, such as applied writing, literature, TV & movie, internet article, historical biography, newspaper and magazine, and translated work.

4.2.2.1 Comparability

To make a corpus-based cross-linguistic analysis, the ideal situation is to use two corpora with the same, or at least comparable, size and make-up. However, to the best of my knowledge there is no sizeable Chinese corpus compiled on the same or similar principles of the BOE. The CCL corpus is the largest general Chinese corpus, which is of roughly comparable size (307 million) with the BOE (455 million). As stated earlier, a large corpus is needed for lexical studies to produce a decent collocational profile, especially for the exploration of less frequent words. Small corpora are likely to show skewing, with too few examples of infrequent words (Moon, 2010) . My thesis will look at most, if not all, of the sad expressions in English and Chinese, some of which are not highly frequent. To obtain a decent collocational profile for them, I would give the priority to the size of the corpus. Since absolute comparability cannot be achieved, what we can do is to find the closest to the one we want to compare it with.

4.2.2.2 Representativeness

Linked to the issue of comparability is the issue of representativeness. Kennedy (1998) has pointed out that representativeness is a matter of judgment and ‘can only be

approximate' (p. 62). Teubert & Cermakova (2004) cast doubts on the notion of representativeness by questioning what a corpus is to represent. They argue as follows:

In any case, we are only justified in claiming that a given corpus is representative of a discourse, however we have defined it, if we have, at least in principle, access to all the texts the discourse consists of. Only then will we have all the relevant information concerning the parameters mentioned above, and only then can we be sure that the corpus we compile as a sample of this discourse is representative...But if this utopia came true, we could do without the corpus. We would already have the discourse as a whole and would not sample it...This is why it does not make much sense to talk about representativeness (p. 117).

It seems that a corpus is never too big if it can be handled computationally. Even a written corpus of 30 million words is small as compared with the actual written texts produced from which it is sampled (Summers, 1991).

The BOE and the CCL corpus are both the largest general corpora available for their own language. Although they were not built based on the exactly same principles, I can assume they are so far the best which can represent English and Chinese, respectively. Since the criteria to build a general corpus are still debatable (Hunston, 2002; Kennedy, 1998; Teubert, 2004a etc.), we cannot say which one is better than the other one because they were both designed following their own principles. My research is to examine how sadness expressions are used and interpreted in both languages, what they share and how they differ, and how correspondences between them can be established. I will not discuss their distributions or percentages contrastively since the data come from two corpora not exactly comparable in composition. Instead, I will first look at monolingual data and present semantic profiles within each language, and then compare them in terms of collocations and semantic associations across languages.

4.2.3 The Babel English-Chinese Parallel Corpus of Fiction

The Babel English-Chinese Parallel Corpus of Fiction (BPC, henceforth) is the fiction part of the Babel English-Chinese Parallel Corpus built at the Institute of Computational Linguistics at Peking University by Chang Baobao et al. I also searched for the emotion words I will examine for this research in Babel's non-fiction part and got nothing because it does not talk about emotions at all. To ease the searches, I will only use its fiction part for my research and refer to it as the Babel Parallel Corpus henceforth. The Babel Parallel Corpus contains a total of 3,707,397 tokens (1,605,850 English words and 2,101,547 Chinese characters). The texts included in the corpus were produced from the 19th century on. All the texts have been aligned at sentence level and all the Chinese texts have been segmented.

4.3 Dictionaries used

A reliance on corpus data does not exclude the use of dictionaries for gathering lexical evidence; rather they are complementary to each other (Ooi, 1998; Teubert, 2000). I will base my analysis on monolingual and parallel corpus data as well as monolingual and bilingual dictionary glosses. The dictionaries to be used in my research include:

Chinese:

("Hanyu Da Zidian," 1993). Wuhan/Chengdu: Hubei Cishu Chubanshe/Sichuan Cishu Chubanshe.

("Xiandai Hanyu Cidian: Bilingual Edition (A Dictionary of Modern Chinese)," 2002). Beijing: Waiyu Jiaoxue Yu Yanjiu Chubanshe.

("Xiandai Hanyu Cidian (A Dictionary of Modern Chinese)," 2005). Beijing: Shangwu Yinshu Guan.

English:

("Collins COBUILD Advanced Learner's English Dictionary," 2006). 5th Edition. London/Glasgow: HarperCollins.

("Oxford Advanced Learner's English-Chinese Dictionary," 2002). (4th edition). Beijing: Shangwu Yinshu Guan).

("Longman Dictionary of Contemporary English," 2006). (4th editon). Harlow, Essex: Pearson/Longman.

("Oxford English Dictionary Online," 2011). Oxford: Oxford University Press.

4.4 Software

4.4.1 The WordSmith Tools

This section introduces one of the two software programs I will use for my research, i.e. the WordSmith Tools (Version 5.0, available at <http://www.lexically.net/wordsmith/>, Wordsmith henceforth), which is probably the most extensively used software for corpus-linguistic analysis. WordSmith has three major functions in, i.e. Concord, WordList and Keyword. My research will mainly use Concord, so the remainder of this section will only introduce this function. Concord does single as well as multi-word searches. It can generate a concordance for the search term, and it also provides the possibility of re-sorting the concordance lines according to the alphabetical order for languages like English or a certain defined order for languages such as Chinese. A concordance is 'a collection of the occurrences of a word-form, each in its own textual environment' (Sinclair, 1991: 32). It is in the form of a concordance that the computer shows us repeated patterns to the right or left of the search term. Another feature of Concord is to generate collocates, clusters and patterns for the search term. Collocation, in Sinclair's words, is the occurrence of two or more words within a short space of each other in a text' and '[a] word which occurs in close proximity to a word under investigation is called a collocate of it' (1991: 170). 'A cluster is a *group of words which follow each other in a text*. The term *phrase* is not used here because it has technical

senses in linguistics which would imply a grammatical relation between the words in it’ (Scott, 2009: 289). In other words, clusters may be found repeatedly in each other’s company, but they are not necessarily meaningful in syntactic sense like patterns, ‘though clusters often do match phrases or idioms’ (p.289). The number of collocates Concord can produce depends on the value settings, such as span and minimum frequency of the word. WordSmith’s default value of collocates is set at 5 to the left and 5 to the right, but users may adjust the horizon up to 25 on either side to meet their own specific requirements.

WordSmith 5.0 works well for many languages, including Chinese. Of course, Chinese needs to be segmented before it can be processed by it. We will talk about this later. Below is the collocate display for Chinese 伤心 generated by WordSmith 5.0:

Figure 4. 1 A screenshot of the collocates of 伤心 by WordSmith 5.0

N	Word	With	l	ation	Texts	Total	al	Left	Right	L5	L4	L3	L2	L1	entre	R1	R2	R3	R4	R5
60	感到	伤心	0.000	1	66	59	7	3	7	10	18	21	0	2	2	3	0	0		
61	但	伤心	0.000	1	65	14	51	1	1	6	2	4	0	28	9	8	2	4		
62	一个	伤心	0.000	1	64	45	19	11	8	9	6	11	0	2	3	3	6	5		
63	死	伤心	0.000	1	63	26	37	9	5	6	4	2	0	11	5	7	7	7		
64	把	伤心	0.000	1	63	26	37	7	4	10	2	3	0	1	17	7	6	6		
65	泪	伤心	0.000	1	63	6	57	1	2	1	1	1	0	16	19	5	12	5		
66	叫	伤心	0.000	1	62	52	10	0	4	13	34	1	0	0	5	3	0	2		
67	母亲	伤心	0.000	1	62	50	12	6	11	7	8	18	0	1	4	2	3	2		
68	去	伤心	0.000	1	61	30	31	9	10	4	2	5	0	1	3	7	11	9		
69	眼泪	伤心	0.000	1	60	7	53	3	2	0	1	1	0	5	24	8	10	6		
70	话	伤心	0.000	1	59	33	26	13	7	6	6	1	0	6	3	4	6	7		
71	看	伤心	0.000	1	59	38	21	4	8	16	8	2	0	2	7	4	4	4		
72	太	伤心	0.000	1	58	49	9	3	1	6	3	36	0	1	1	1	1	5		
73	对	伤心	0.000	1	57	16	41	6	2	3	5	0	0	4	12	6	15	4		
74	别	伤心	0.000	1	57	48	9	1	2	9	10	26	0	1	6	1	1	0		
75	大	伤心	0.000	1	56	24	32	9	4	5	3	3	0	3	8	7	10	4		
76	却	伤心	0.000	1	56	31	25	9	9	6	5	2	0	9	2	7	4	3		
77	现代	伤心	0.000	1	54	2	52	2	0	0	0	0	0	0	23	10	10	9		
78	什么	伤心	0.000	1	54	30	24	5	4	5	6	10	0	8	4	3	4	5		
79	他们	伤心	0.000	1	52	29	23	1	4	3	8	13	0	2	4	8	4	5		
80	读者	伤心	0.000	1	52	1	51	0	1	0	0	0	0	0	0	0	17	34		
81	个	伤心	0.000	1	52	30	22	12	6	4	3	5	0	0	2	5	5	10		
82	处	伤心	0.000	1	51	4	47	0	2	0	1	1	0	37	5	1	3	1		
83	知道	伤心	0.000	1	51	29	22	10	7	7	3	2	0	1	6	3	8	4		
84	落泪	伤心	0.000	1	50	3	47	0	1	0	1	1	0	42	4	1	0	0		
85	没	伤心	0.000	1	50	20	30	7	8	3	1	1	0	2	8	5	6	9		
86	再	伤心	0.000	1	49	23	26	7	3	6	1	6	0	2	3	6	8	7		

4.4.2 ParaConc

ParaConc is a parallel concordance program via which recurrent patterns in translated texts can be identified. ParaConc can handle up to four parallel texts, which could be an original text with three different translations or four different languages. The translated texts have to be aligned before they can be analysed by ParaConc because no ‘language-particular information is encoded within the program’ (Barlow, 2008: p 106) and the search is done in a mechanical manner. ‘Alignment is the provision of information on equivalent text segments in the two languages, typically using the sentence unit as the basic alignment segment’ (Barlow, 2008: 106). Usually, the source and translated texts in the aligned parallel corpus are put in separate files, with the source and target text of each sentence being linked together with a unique identifier. The following sample shows what the English text looks like after the alignment:

```
<p id='3'>
_<a id='3' no='1'>
  <s id='1'> On his bench in Madison Square Soapy moved uneasily. </s>
</a>
_<a id='4' no='1'>
  <s id='2'>When wild goose honk high of nights, and when women without
    sealskin coats grow kind to their husbands, and when Soapy moves uneasily on
    his bench in the park, you may know that winter is near at hand.</s>
</a>
</p>
```

The sentence number is inserted as the value of an ‘id’ attribute of the <s> tag. The ‘id’ is organised in such a way that it will be a unique identifier to link the sentences in Chinese and English:

```
<p id='3'> paragraph 3
  <a id='3' no='1'> alignment pair 3 (consisting of 1 sentence),
  <s id='1'> sentence 1
```

The id number for the alignment pair is unique as it is counted throughout the text.

Alignment is an essential step in the construction and exploitation of parallel corpora.

Its aim is to make it possible to retrieve mechanically corresponding texts from the original and the translation by using a search program. To show how an alignment is

done, a sample of aligned texts by using Paligner, developed by the Institute of

Computational Linguistics at Peking University, is given below:

(1) ●●

[[1:1]]

<s>老师已经表示过，就算能把它弄到车上，到了他要去的基督堂[注] 那个城市，他还是不知道拿它怎么办，因为他初来乍到，只能临时找个地方住住。 </s>

<s>The master had remarked that even if he got it into the cart he should not know what to do with it on his arrival at Christminster, the city he was bound for, since he was only going into temporary lodgings just at first.</s>

(2) ●●

[[2:1]]

<s>一个十一岁的男孩子正帮着扎东西，挺有心事的样子，这时走到大人这边来，趁他们摸着下巴颏的时候，大声说：‘姑婆有个好大的柴房哪，你找到地方放它之前，也许能寄放在那里头吧。’</s>

<s>他因为说话声音大，脸红了。 </s>

<s>A little boy of eleven, who had been thoughtfully assisting in the packing, joined the group of men, and as they rubbed their chins he spoke up, blushing at the sound of his own voice: ‘Aunt have got a great fuel-house, and it could be put there, perhaps, till you've found a place to settle in, sir.’</s>

(3) ●●

[[1:2]]

<s>裘德呀，你这会儿还不懂我走的道理，等你再大点，你就明白啦。’</s>

<s>You wouldn't understand my reasons, Jude. </s>

<s>You will, perhaps, when you are older.’</s>

The markup illustrated here is all inserted automatically when the program is run. The

markers ●●, used here identify paragraphs and alignment pairs respectively. Angle

brackets are used for tags: <> stands for ‘start-tags’ and </> for ‘end-tags’. <s> ‘You

will, perhaps, when you are older.’</s> means that this is a single sentence in its

original text. [[1:2]] in example (3) shows that a single sentence corresponds to two sentences in the other language. There are also cases where two sentences are aligned with a single in the other language marked by [[2:1]], as illustrated in example (2). In example (3), the Chinese sentence ‘裘德呀，你这会儿还不懂我走的道理，等你再大点，你就明白啦’ was split into two sentences in English. Texts with quotations are prone to cause errors, as shown below:

- (4) [[1:1]]
<s>‘我可没那么多钱买。’</s>
<s>‘I cannot afford that,’ said Sue. </s>
[[2:1]]
<s>她说。</s>
<s>她还的价非常之低，再没想到，卖像人居然把拴像的细铜丝解开，隔着篱梯把它们递过来。</s>
<s> She offered considerably less, and to her surprise the image-man drew them from their wire stay and handed them over the stile. </s>

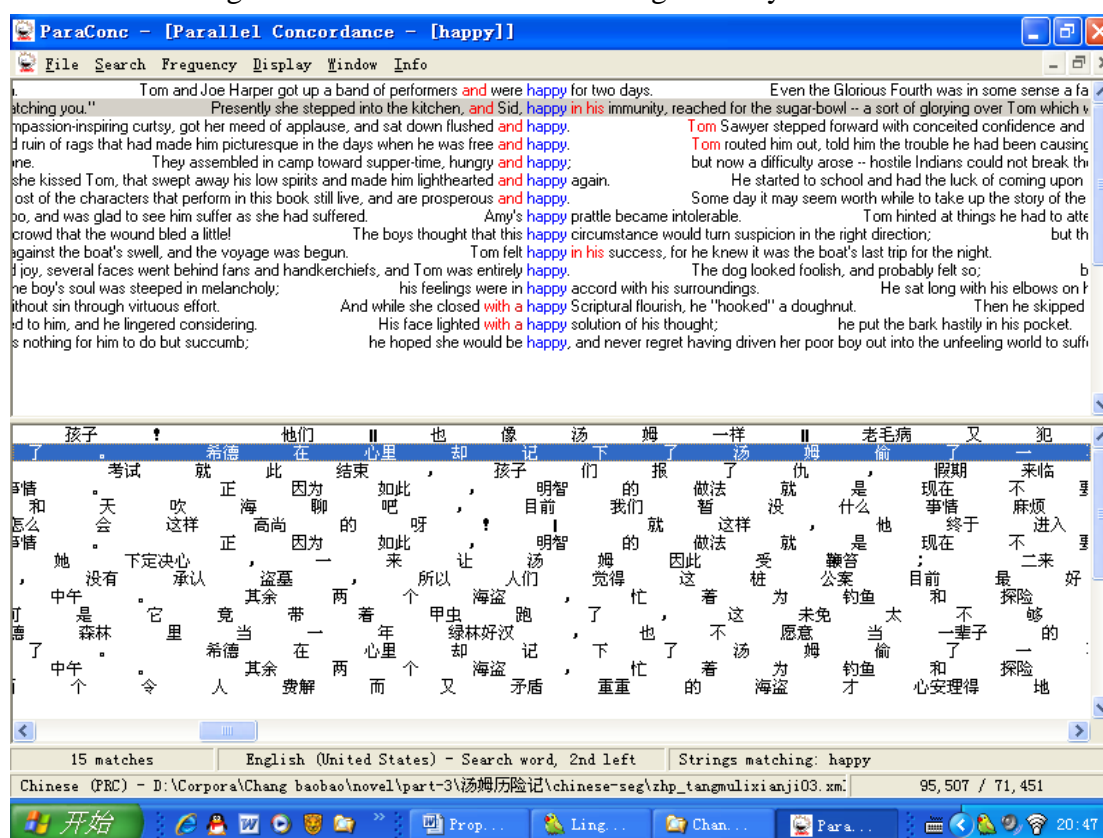
The misalignment in example (4) occurs because 她说 (*said Sue*) is followed by a stop 。, which is a comma in its original text. Of course, the error here is trivial, but sometimes it can lead to a mess. The approach used in this program is a statistical one. It is based on sentence length in terms of words and characters per sentence in Chinese and English.

Compiling a parallel corpus is much more difficult than a monolingual one. Firstly, all automatically aligned texts, however well the program works, need to be manually post-edited to check errors. The post-editing work is usually quite heavy, making alignment a time-consuming task. Besides, the compilation of parallel corpora will be greatly restricted by the availability of translated texts, since the number of source texts which have been translated into another language is limited. Their lesser

availability in machine-readable form will make this task even more difficult to accomplish. In addition, there is another issue to consider: the quality of translations, which relates closely to the reliability of the data retrieved from the corpus, as well as to the progress of alignment. Therefore, it seems improper to use the same criterion of representativeness for monolingual corpora to judge parallel corpora.

What it looks like after running ParaConc is shown in Figure 4.2:

Figure 4.2 A screenshot of searching result by ParaConc



ParaConc presents the user with '(i) multiple instances of the search term and (ii) a large context for each instance of the search term, thereby allowing a thorough analysis of usage in terms of the equivalences between two languages' (Barlow, 2008: 105). The concordance lines can be sorted in the alphabetical order for languages such as English or a certain order defined for other languages such as Chinese (See Miachel Barlow's

homepage for how to use the software: <http://www.michaelbarlow.com/>). By using this program, we can obtain some empirical evidence on ‘the congruence in particular forms between two languages’ (Barlow). For example, drawing on the parallel data retrieved from a bidirectional English-Chinese parallel corpus, we can examine the congruence between the English word *sorrow* and the Chinese word *beishang*. How likely is that *sorrow* will be translated by *beishang* and how likely *beishang* will be translated by *sorrow*? For more polysemous words such as *unhappy*, a more detailed investigation will be needed to examine the different senses and translations of *unhappy* to identify what Chinese equivalent expression(s) each sense is likely to correspond to. The congruence issue across languages has been greatly complicated by the association of more than one meaning with the same linguistic form, which is inescapable in all human languages.

ParaConc provides a variety of frequency statistics, of which the most relevant kind to my investigation is collocate frequency. By clicking on the COLLOCATE FREQUENCY DATA it gives the collocates of the search term ranked according to frequency. It offers collocates within the span of -4 to 4. It is difficult for the program to locate translations automatically. The Hot Words function may suggest some possible translations or associated words such as collocates of the search term, but it does not work well for a less morphological language like Chinese. In an investigation concerning Chinese, we have to examine the concordance lines manually to find the candidates of equivalence. Based on their frequency, translation equivalents can be

identified. Therefore, all I need from ParaConc for my research is the parallel concordance lines and frequency data of the search term.

The advantage of using parallel data to obtain information on congruence between language pairs is that we do not rely on intuitions, but objective and quantitative information on correspondences, to judge congruence. However, we have to bear in mind that the parallel data will vary according to the corpus used in the investigation, and that the interpretation of corpus data will still involve researchers' intuitions about language, though it is conducted on the basis of quantitative data (Barlow, 2008).

In this section I have briefly introduced a parallel concordancer, *ParaConc*, which is highly useful for contrastive analyses. I have examined some of its important features and the rationale of the program. The frequency data as well as the parallel texts generated by this program can guide our investigations into similarities and differences of lexis or constructions between compared language pairs (Barlow, 2008).

4.5 Data processing and analysis

4.5.1 Clean-text policy and annotation

Views differ as to whether corpora should be kept unprocessed or annotated with part of speech (POS) or other codes. Sinclair strongly favours 'the clean-text policy' for two reasons (1991: 21):

Firstly, each particular investigation is likely to view the language according to different priorities. Its analytic apparatus may well be valuable and interesting to the next investigator, and even adaptable to the new needs; but not so standardised that it can become an integral part of the corpus. Secondly, although linguists leap effortlessly to abstractions

like ‘word’ (meaning lemma, see Chapter 3) and beyond, they do not all leap in the same way, and they do not devise precise rules for the abstracting. Hence, even the bedrock of assumptions of linguistics, like the identification of words, assignment of morphological division, and primary word class, are not at all standardised. Each study helps the others, but does not provide a platform on which the others can directly build.

Sinclair thinks the analytical systems derived from traditions of language analysis are problematic, even the traditionally established parts of speech, so we should reject them all and derive some new categories from raw corpora. We have discussed this issue in Section 3.1. Contrary to Sinclair, Leech (1997) argues that annotation can enrich the corpus and therefore facilitate investigations. This opposition is also related to the distinction between corpus-based and corpus-driven approaches to language studies. Sinclair’s idea is ‘to refrain from imposing analytical categories from the outside until we have had a chance to look very closely at the physical evidence’ (1991: 29). It seems that annotation should be used with caution before we are sure about the analytical categories we will assign to the texts. I will talk about this issue when I discuss Chinese POS tagging in Section 4.5.2.2.2.

4.5.2 Data processing

4.5.2.1 English data

For the English data in this investigation, I will draw upon the BOE, which allows the investigator to expand concordance lines to five lines of co-text, so it is possible to examine the wider contexts. It is POS tagged and available online at <http://wordbanks.harpercollins.co.uk/>. The BOE is accessed and manipulated via Sketch Engine (<http://www.sketchengine.co.uk>). Sketch Engine is a software program which was developed by Lexical Computing Ltd. More detailed information can be found in

Kilgarriff & Tugwell (2001), Kilgarriff & Rundell (2002), Kilgarriff, Rychly, Smrz, & Tugwell (2004), Krek & Kilgarriff (2006) and Kilgarriff (2009) etc. The interface for its concordancing options is shown below:

Figure 4.3 The interface of the concordancing options of word sketch

The screenshot shows the 'Word Sketch' interface with the following elements:

- Navigation tabs: Home, Concordance, Word List, **Word Sketch**, Thesaurus, Sketch-Diff
- Corpus: WordBanks Online: English
- Query: [Empty text box]
- Keyword(s): [Expandable section]
- Context: [Collapseable section]
- Query Type: All of these items.
- Window Size: 5 tokens (Left context) and 5 tokens (Right context)
- Lemma: [Empty text boxes for Left and Right context]
- PoS: (use Ctrl+click for multiple selection)
 - Left context: adjective, adverb, conjunction, determiner
 - Right context: adjective, adverb, conjunction, determiner
- Text Types: [Expandable section]
- Buttons: Make Concordance, Clear All

As we can see from the above figure, Sketch Engine produces concordances, word lists, thesaurus, word sketches and sketch differences. Word Sketch provides collocates of a given word occurring in its main grammatical relations/patterns according to frequency or statistical salience. The present statistics used is logDice, based on Dice coefficient (see Rychly, 2008). Sketch Diff compares the grammatical and collocational behaviour of two semantically related words and show what they share and how they differ. Sketch Engine works on lemmatised forms, so all of the grammatical forms of a given word would be considered unless it is set to search for only one form. All collocate searches will be made on the basis of lemma and the span will be set at 5 to the left and 5 to the right for word sketches and sketch differences, and -1 to +1 for top 50 collocates, unless otherwise stated. In terms of statistic measurement, the top 50

collocates will be based on t-score and the word sketches and sketch differences on salience.

4.5.2.2 Chinese data

4.5.2.2.1 Chinese data processing

Unlike English or other western languages, Chinese does not delimit words by space, so word segmentation is prerequisite to Chinese language processing. Chinese words are formed by combination of characters and one character may occur in any position of a word, so segmentation ambiguities may occur in a text, as the following examples show

(Goh, Asahara, & Matsumoto, 2005: 382)

不/可以/淡忘/远在/故乡/的/父母/
not/ can/ forget/ far away/ hometown/ DE/ parents/
(Cannot forget parents who are far away at home)
不可/以/营利/为/目的/
cannot/ by/ profit/ be/ intention
(Cannot have the intention to make a profit)

As shown above, the character 可 (*ke*) can combine with 以 (*yi*) to form 可以 (*keyi* /can) or combine with 不 (*bu*) to form 不可 (*buke/cannot*). Segmentation is done based on the context in each sentence. Many researchers in computational linguistics have been developing various systems to improve the segmentation performance in order to achieve a high accuracy rate.

The CCL corpus will be the main source of my Chinese data. It is a raw corpus, unsegmented and POS untagged. Besides, its online concordancer generates only the concordance of the search item, but no collocates. So I will download all the concordance lines containing the search item and segment them by using FreeICTCLAS (Institute of Computing Technology, Chinese Lexical Analysis System)

developed by the Institute of Computing Technology, Chinese Academy of Sciences (for details see <http://www.ict.ac.cn/technic/channel/detail3432.asp>). According to its official website, its segmentation precision rate is 97.58%. Finally, the segmented texts/lines will be processed by using WordSmith 5.0 to obtain both collocates and concordances. For the collocates, the span will be set at 5 to the left and 5 to the right.

4.5.2.2.2 Chinese part-of-speech classification and tagging

Concepts are encoded as words in language. A given concept may not relate to the same word class in different languages. The classifying criteria for word classes might be particular to each language (Dixon & Aikhenvald, 2004). The recognition of word classes in an inflectional language is purely based on its syntactic distribution. For Chinese, there is no simple one-to-one correspondence between word class and syntactic function (Zhu, 1985: 4) because it has little inflectional morphology. Since *Mashiwentong* (马氏文通), the issue of word class classification has been a tricky problem in the field of Chinese grammar, where no agreement about it has been made (M. Hu, 1996). The main reason is due to the fact that a Chinese word with the same meaning may perform different syntactic functions in different contexts. Consider the following examples containing *beishang*:

[悲伤]的泪水	‘sad tears’
装满了个人的烦恼和[悲伤]	‘be filled with private worries and sadness’
[悲伤]地哭了	‘cried sadly’

The English translations above are all direct ones, from which it is clearly seen that *beishang* can function as an adjective, noun or adverb, meaning almost the same. It has

been generally agreed that each Chinese word has two properties: semantic property (lexical meaning) and grammatical property (function). The word class recognition in Chinese can be based on meaning or function or both. Syntactic function is not the indisputable criterion in Chinese word classification, and the fuzzy concept of meaning is involved in, so there is wide disagreement about this issue. Most mainland Chinese scholars focus on meaning or both, while non-mainland scholars, including those from Taiwan, tend to base their classification on syntactic function alone. In the mainland, the main-stream point of view is that an adjective, in some cases, can be used as a noun or adverb, but if this happens not very frequently, then it is better to say it is only used as another word class, not that it has acquired one more word class membership (Y. Hu, 1995 etc.). It is still an adjective and just sometimes used as a noun. Such a phenomenon is referred to as *huoyong* (活用). This idea has originated from Lú Shuxiang (1954, 1984) who proposes that a word should belong to a fixed category and the proportion of multi-category words should be minimised. This idea has become a guiding principle in classifying word classes in Chinese and been widely accepted by Chinese linguists. Recently, however, it has been challenged by some researchers, such as (Gao, 2009), who looks at Chinese word classes from a cognitive perspective and claims that since word classes are symbolic structures that emerge from language use, the proportion of multi-category words should be determined by the empirical data, rather than adhering to a priori principle. He argues that Lú's theory is an idealised one which does not conform to the Chinese linguistic facts. He gives some examples, in which the

underlined words only marked as adjectives in *Xiandai Hanyu Cidian* (5th edition, the 5th

XHC henceforth) are used as nouns (p. 21):

自己富了,也得设法为国家分担忧愁(*youchou*)。
他声音里不但充满骄傲(*jiacao*),也充满自信(*zixin*)。
很多人对开刀都很害怕(*haipa*),有的在做一些小手术时都会产生恐惧(*kongju*)。
酒虽然不能解除任何痛苦(*tongku*),至少总可以使人暂时忘记。
基因工程的发展对人类的贡献是无法估量的,但也带来了一些烦恼(*fannaoyao*)。
后来,梁大爷虽然脱离了危险(*weixian*),但慢慢地瘫痪在床。

Many of the underlined words above are emotion words, such as 忧愁/worries, 恐惧/fear, 痛苦/anguish and 烦恼/vexation. In fact, such phenomena are not uncommon in Chinese. Ma (1994) makes a statistical analysis in a one-million word corpus and finds that many emotion words, such as 烦恼/vexation, 欢乐/joy, 苦恼/distress, 快乐/gaiety, 痛苦/anguish, 喜悦/rejoicing and 忧虑/anxiety, are used more frequently as nouns than as adjectives, while for others, such as 悲哀/mournful (13:4), 悲伤/sorrowful (8:3), 悲痛/grieved (10:4), 愤怒/indignant (11:4) and 羞愧/shamed (7:3), the adjectival use still dominates the picture. He proposes to classify a word according to its main use on a statistical basis. But obviously we cannot be convinced that 痛苦 is a noun but 悲伤 is an adjective since they all denote feelings.

What makes the issue more complicated is that there is a tradition of saying Chinese does not distinguish verbs and adjectives (Bhat, 1994; Schachter, 1985; Thompson, 1988)⁶. Quite a few linguists argue that Chinese adjectives (at least a subset

⁶ Schachter uses ‘nage nūhaizi liaojie/piaoliang’ and ‘liaojie/piaoliang de nūhaizi’ to illustrate ‘there is no consistent basis for distinguishing verbs with adjectival meanings from other verbs’ (p.18). But, in fact, in the above examples, *piaoliang*/beautiful and *liaojie*/understand are not the same. For one thing, ‘nage nūhaizi liaojie’ (that girl understands) has to be embedded in the context to be understood, otherwise it is incomplete in meaning, while ‘nage nūhaizi piaoliang’ (that girl is beautiful) is a perfect and independent

of it) have some grammatical properties in common with verbs. It is really a subtle matter to distinguish verbs and adjectives based on syntactic function alone. Consider the following examples from Zhu Dexi (1980):

今儿怪冷的。
价钱很便宜。
他心里空空洞洞的，什么也不怕。
浑身软绵绵的...

In the above examples, the underlined words are all adjectives, which function as intransitive predicate in the above sentences. They are similar to verbs in terms of grammatical property. There is another set of adjectives in Chinese, like *jin* (gold), *yin* (silver), *nan* (male), *nǚ* (female), *manxing* (chronic), and *daxing* (large-scale), which can only function as attributive modifiers. They are referred to as ‘non-predicative adjective’ or ‘determiner’ (Zhu, 1982). Some scholars, (e.g. J. Li, 1924; Shen, 1999; L. Wang, 1955; B. Zhang & Fang, 1996; Z. Zhang, 1979) argue that the main function of Chinese adjectives is to serve as attributive modifiers. However, some other scholars, such as Chao (1968), Li & Thompson (1981), Zhu (1982), view adjectives as a sub-class of verbs, on the basis of the belief that Chinese adjectives mainly function as predicates. McCawley’s (1992) investigation further supports this viewpoint. He examines Chinese adjectives by using the distinctive features between verbs and adjectives in many other languages and comes to a conclusion that Chinese adjectives are exactly verbs, rather than a subset of it with adjectival properties. From the above discussion, it can be seen that Chinese adjectives are quite different in grammatical properties from the familiar

sentence. For another, ‘liaojie de nǚhaizi’ (an understanding girl) is ungrammatical, while ‘piaoliang de nǚhaizi’ (a beautiful girl) is standard Chinese.

kind of adjectives in English. The debate on their scopes and functions is still going on. I do not see it ending soon.

It is well-known that linguistic applications have to depend on the development of linguistic theories. The theoretical disagreement about Chinese word class classification has seriously hindered its application to Chinese lexicography. The great majority of old Chinese dictionaries simply evaded this issue by not marking the word class. The 5th XHC has made an attempt to touch on this issue, but as mentioned earlier, the result is unsatisfactory.

The controversy over the Chinese word class classification has also posed a problem for automatic POS tagging in corpus linguistics. Since there have been no well-accepted criteria for recognizing word classes in Chinese, different POS tagging tools may adopt different tagsets to accomplish their tasks. Xia (2000: 4) makes the following comments on this issue:

The central issue in Part-of-Speech (POS) tagging is whether the POS tagging should be based on meaning or on syntactic distribution. This issue has been debated since the 1950s (Gong, 1997)⁷ and there are still two different viewpoints. For example, a word such as 毁灭 in Chinese can be translated into destroy/destroys/destroyed/destroying/destruction in English and it is used roughly the same way as its counterparts in English. According to the first view, POS tags should be based solely on meaning. Because the meaning of the word remains roughly the same across all of these usages, it should always be tagged as a verb. The second view says that POS tags should be determined by the syntactic distribution of the word. When it is the head of a noun phrase, it should be tagged as a noun in that context; when it is the head of a verb phrase, it should be tagged as a verb.

⁷ The original [Gon97] is changed to (Gong, 1997) to make it consistent with the citation format in this thesis.

Most projects outside the mainland of China, such as Sinica Corpus (Academia Sinica Balanced Corpus of Mandarin Chinese, in Taiwan) (http://rocling.iis.sinica.edu.tw/CKIP/paper/wordsegment_standard.pdf) and CTB (Penn Chinese Treebank), choose syntactic distribution as the main criterion for their POS tagging because it ‘complies with the principles adopted in contemporary linguistics theories, such as the notion of head projections in the X-bar theory and the GB theory’ (p. 5). Some linguists tend to use western linguistic theories to explain all Chinese linguistic facts. For some phenomena, it is true that they do; however, for some others, they do not work well. For example, 政治/*politics* is a noun, but it is tagged as an adverb in ‘只[only] 能[can] 政治[politically]/AD 解决[resolve]/VV 这个[this] /M 问题[problem]’ (This problem can only be resolved through political means) in CTB (Xia, 2000) . It is hard to accept that 政治 has become an adverb just because it is used in a sentence with omission. Look at another example:

去 行, 不 去 也 行
 Qu xing bu qu ye xing
 Go ok not go also ok

If solely based on function as *politically* was treated in that sentence, it seems that *qu/go* in this example should be tagged as a noun. However, few Chinese scholars would agree that *qu* is a noun.

ICTCLAS bases its tagging on both meaning and function. For example, it divides adjectives into four subclasses: ad/副形词(both adverb and adjective), an/名形词(both adjective and noun), ag/形容词性语素(adjectival morpheme) and al/形容词性惯用语(adjectival idioms). Taking meaning into consideration, ICTCLAS is rather

conservative about POS tagging, trying to achieve the highest accuracy and not to overgeneralise about POS classification. For example, the team makes reference to CTB, and assign specific tags to ‘是’/be, but they do not use the tag of ‘copula’ because they are aware that ‘是’ has many functions, of which copula is only one. It is extremely difficult to distinguish these functions on the basis of the present technology. If you do it reluctantly, it will lead to a decline of accuracy. To ensure a high accuracy, they choose not to go further now (http://www.ictclas.org/hottopic_010.html). Of course, the result of conservatism is that emotion words, like *beishang*, *kuai*, are marked as *an*, which means they can be used as both an adjective and noun. The adverbial use is not considered, probably because they do not think they are also adverbs, but just *huoyong* (used) as adverbs sometimes. The disambiguation of multi-category words is one of the difficulties in part-of-speech tagging. Generally speaking, Chinese words are more difficult to disambiguate in terms of POS than English (Huang, Harper, & Wang, 2007: 1093). Tseng, Jurafsky, & Manning (2005) observed that, in comparison to 19.8% of the English words in WSJ (Wall Street Journal), 29.9% of the words in CTB have more than one POS possibilities. According to Huang et al. (2007), the highest accuracy for English has reached 97% using various models (Brants, 2000; Ratnaparkhi, 1996; Thede & Harper, 1999), while their re-ranking model using n-gram and morphological features for Chinese has only achieved an accuracy of 94.16%.

As noted earlier, the Chinese corpus I am using for this research is not POS tagged. Currently, the only publically available POS tagging tool is ICTCLAS. Since it does not distinguish adjectival, nominal and adverbial uses of emotion words, there is no

point in POS tagging my Chinese data in this study. Besides, Sinclair's clean-text policy suggests that it is better to defer the use of POS categories to 'refrain from imposing analytical categories from the outside' (1991: 29). Analysing untagged Chinese data will lead to less correspondence between the methodological details for English and Chinese. For English, I will search for all the relevant forms of the same concept separately, for example, *sorrow*, *sorrowful*, *sorrowfully*, because they have totally distinct collocates. For Chinese, I will search for only one form, e.g. *beishang*. In doing so, one main disadvantage would be, for the Chinese sadness expressions, I will not be able to present a detailed quantitative analysis of their colligational patterns. Instead, I will focus on some main colligational features and compare them with their English counterparts. Due to the constraints on the development of Chinese POS classification and therefore the tagging technology, I will compare them as a whole across languages, whether as an adjective, noun or verb, or whether functioning as attribute, subject, object, predicate or adverbial. Since the focus of my research is on the exploration of sadness concepts across cultures, the research findings will not be considerably affected.

4.5.3 Data analysis

For each group of words, the analysis will start with the discussion of the English expressions. Before analysing the corpus data, their definitions from the three corpus-based dictionaries will be examined: *Collins COBUILD Advanced Learner's English Dictionary* (COBUILD, henceforth), *Oxford Advanced Learner's English -Chinese Dictionary* (OALECD, henceforth) and *Longman Dictionary of Contemporary English*

(LDCE , henceforth). The *Oxford English Dictionary Online* (OED, henceforth) will also be consulted when more information is needed. For corpus data, I will seek to figure out their colligations and semantic associations/preferences by examining their collocates. Collocates can be identified automatically by the software. Based on collocates, colligations are not difficult to identify since they are quite objective. In English, as the members of the same derivational family have quite distinct colligational properties, especially in terms of prepositions, I will list them separately. Semantic preference/association is mainly concerned with tendency and probability. In this research, I will focus only on emotion-related collocates and classify them into various groups and assign them labels to indicate their semantic preferences/associations. The establishment of the lexical links across languages will be made mainly on the basis of the data obtained from the parallel corpus. For Chinese, as I do not distinguish word classes for sadness expressions, I will deal with them together. As Chinese sadness expressions are very similar in their colligational patterns, I will only outline them by highlighting the differences across languages in the summary of the analysis (Section 5.4). The significant collocates will be categorised according to *introversion/extroversion*, *modifier*, *feeling* and *others*, which include the words helpful for uncovering the meaning. The collocates will be selected on the basis of raw frequency because the calculation of t-score by Wordsmith requires the whole corpus. What I have for Wordsmith to process is only the concordance lines extracted from the CCL corpus. Obviously, my methodology for this research is greatly limited by the availability of corpora and the current technology. Nevertheless, the selection of Chinese collocates

according to the raw frequency will not significantly affect my findings because I will not compare the collocate frequencies of semantically similar expressions across corpora because they are statistically incomparable. For example, in Chinese, *ganjue*, *gandao*, *juede*, and *jue* are frequent collocates of *nanguo*, all meaning 'feel'. If we make a quantitative comparison of their frequencies or t-scores or whatever between *nanguo* and its English equivalent(s), should they be counted as one word or four in actual counting or statistic calculation? It appears that a precise or statistical comparison across languages is problematic before we provide an answer to the question of equivalence between languages. Basically, I will rank their collocates separately, present the significant ones based on t-score or raw frequency and obtain a collocational profile of their own. I may compare synonyms based on the same measurement within the same language. The rationale of the analysis is to examine a lexical item or compare synonyms within each language and tease out the distinctiveness of their own and then contrast it/them qualitatively with its/their equivalents in another language. The idea is that by inspecting the collocates we can get abundant semantic information about the examined item. In addition to collocates, I will also look for explicit and implicit paraphrases, i.e. real paraphrases and typical examples, in the corpus. After obtaining the semantic profiles for sadness expressions in the two languages, a cross-linguistic qualitative comparison will be made based on their colligations, collocations, semantic references/associations and other features implied in their collocates and examples/paraphrases. Finally, a summary of the research findings will be presented.

CHAPTER 5

CONTRASTIVE ANALYSIS OF SADNESS EXPRESSIONS IN ENGLISH AND CHINESE

5.1 Overview

Psychologists tend to believe that there are some basic human emotions, which are innate and universal. For instance, Ekman, an extreme universalist, claims that basic emotions like *happiness*, *sadness*, *anger*, *fear*, *disgust* and *surprise* are linguistically recognised in all languages (Ekman, 1973, 1993; 1994 etc.). However, linguists are more likely to favour the view that different languages divide emotions into different categories and put different labels on them. Language-specific emotion words might overlap, but they do not exactly map onto each other across languages. Analysis of the emotion lexicon of a language helps reveal the way in which its speakers cut the emotional world. A cross-linguistic comparison of the emotion lexicon can therefore disclose similarities and differences in the conceptualisation of emotion across cultures.

This research aims to focus on the question whether the concept of *sadness* is universal by comparing and contrasting English sadness expressions and their Chinese counterparts in terms of certain structural categories proposed in the field of corpus linguistics. It bases its claims mainly on colligations, collocations, semantic associations and paraphrases, focusing on identifying what elements are common to both languages and what elements are specific to English or Chinese. Throughout this chapter, I will engage with the corpus-based analysis of English sadness expressions and their closest

counterparts in Chinese. Some of them will be discussed in pairs due to their close similarity in meaning. Each or each pair of the terms is the subject of an individual study, and an exercise in multifaceted cross-cultural analysis, within a corpus-based methodological framework. During the discussion, the cultural underpinnings for language- or cultural-specific categorisations will also be explored. This chapter makes an attempt at a systematic account of English sadness expressions and their Chinese counterparts. I consult different dictionaries and thesauri to produce a list of English emotion expressions to be included in the analysis, and some of them are excluded after careful scrutiny for some reasons, such as low frequency in the corpus. Their Chinese equivalents are determined basically based on the search results in the parallel corpus. Because the parallel corpus used for this research is a uni-directional one, i.e., from English to Chinese, the Chinese list yielded is longer than the English one. While it is not an exhaustive discussion of all the sadness expressions in the two languages, it does cover most of the frequent sadness expressions, totalling about 100 including their derivational forms and Chinese equivalents, which constitute the core of the English and Chinese sadness lexicons.

5.2 Contrastive Analysis of Sadness Expressions in English and Chinese

5.2.1 *Sorrow** & *grief** and their Chinese equivalents

5.2.1.1 *Sorrow** and *grief**

Sorrow and *grief* are two common emotion-denoting terms in English, which are usually viewed as synonyms. However, in the following sentences, *grief* cannot be replaced by *sorrow*, and vice versa (Teubert, 2004a: 82):

- 1) *Grief* gave way to a guilt that gnawed at him.
- 2) A magic harp music made its listeners forget *sorrow*.

So, in a strict sense, they cannot be viewed as synonyms. Native speakers of English may understand their meanings, but they are ‘less competent in describing’ them (Teubert, 2004a: 93). To determine what emotion concepts they are associated with respectively, I turned to the three corpus-based dictionaries: LDCE, OALECD and COBUILD. There I found the following entries, with the idioms and some examples left out (henceforth):

LDCE:

sorrow n

1[U] a feeling of great sadness, usually because someone has died or because something terrible has happened to you 2[C] an event or situation that makes you feel great sadness

grief n

1[U] extreme sadness, especially because someone you love has died 2[C] something that makes you feel extremely sad 3[U] *informal* trouble or problems

OALECD:

sorrow n

1[U] ~ (**at/for/over sth**) feeling of sadness or distress caused esp by loss, disappointment or regret; grief 悲伤; 悲痛; 懊丧; 悔恨 2[C] particular cause of this feeling; misfortune 悲伤的原因; 懊丧的原因; 不幸

grief n

1[U] ~ (**over/at sth**) deep or violent sorrow 忧伤; 悲伤 2[C] event causing such feelings 引起忧伤、悲伤的事; 伤心事

COBUILD:

Sorrow is a feeling of deep sadness or regret.

Grief is a feeling of extreme sadness.

According to these dictionaries, it seems obvious that *sadness* is the hypernym of *sorrow* and *grief*, and *grief* is a more intense feeling of *sadness* than *sorrow*, though both of them can be caused by someone’s death. A close inspection of the definitions above shows that both *sorrow* and *grief* basically refer to a sad feeling or an event/situation that causes such a feeling. The only exception is that *grief* is also used

informally in the sense of ‘trouble or problems’ (LDCE), which is an extension of its second sense ‘something that makes you feel extremely sad’. Since it is not used very frequently as neither OALECD nor COBUILD lists this as a separate sense, we can regard it as peripheral. Our discussion will focus on how to distinguish *sorrow* and *grief* from each other, in the sense of feeling sad. Do they differ only in intensity? Why in the sentences mentioned earlier they cannot be substituted for each other? If we want to know their subtle nuances of meaning and differentiate them from each other, we need to draw on corpus data.

To find the colligational patterns of an English sadness expression, I mainly look at its collocates (examined as lemma) within the span of -1 and +1. I am fully aware that the restriction of collocate search to such a short span will result in some perplexing lacunae, but the extension of the search span will unavoidably lead to some inaccuracies. Besides, I inspect its concordance, especially the words immediately before and after the search word, to filter out those irrelevant collocates. I also use word sketch and my intuition to complement this kind of search to get a fuller semantic profile of the search word. This procedure will be followed for searching for colligation properties throughout the thesis, unless otherwise stated. For the sake of space, I will not present all the tables and offer a step-by-step analysis, but only summarise the search findings in most cases. For *sorrow* and *grief*, the BOE yields 3469 and 8183 hits (as a noun and lemma, retrieved on Sep 1st, 2010⁸) respectively. According to the corpus data, *sorrow* has strong colligations with prepositions, typically *of* (546), *in*

⁸ All the English data in the thesis were updated in September, 2010.

(243), *for* (169), *with* (155), *at* (108), *over* (64), verbs, predominantly *express* (49), *feel* (37), *experience* (16), *bring* (11), *drown* (8), modifying adjectives, principally *great* (100), *deep* (77), *profound* (26), *sweet* (16), *heartfelt* (11), *genuine* (11), *real* (11), *personal* (9) and possessive adjectives, such as *his* (190), *their* (148), *your* (103), *her* (96), *my* (74), *our* (69), *own* (11).

For *grief*, similar colligational patterns are found: prepositions, typically *of* (1352), *with* (490), *in* (434), *for* (289), *over* (155), *to* (418), *at* (189), *about*(57), *into* (48), modifying adjectives, like *good* (262), *private* (70), *great* (44), *personal*(36), *public* (26), *deep* (22), *terrible* (20), *unresolved* (19) and verbs, mainly *express* (35), *cause* (33), *feel* (35), and possessive adjectives, such as *their*(369), *her*(307), *his*(261), *our*(131), *my*(128), *your*(91), *own* (73). 262 hits of *good* are found in its collocates because ‘good *grief*’ is a fixed phrase used ‘when you are slightly surprised or annoyed’. It also frequently collocates with *counselor*(47)/*counsellor*(25)/*counsel*(24), suggesting that in modern society *grief* has increased to be a problem and needs to be handled with professional help.

To find the differences between *sorrow* and *grief* in semantic associations, I looked at their sketch difference. Their collocates in the ‘and/or’ pattern are shown below (the first two groups of figures show the frequencies for *sorrow* and *grief* respectively and the figures in the last two columns indicate the salience, henceforth):

Table 5.1. 1 The common pattern of *sorrow* and *grief* in the BOE: and/or

and/or	1180 s	1971 g	5.6 s	4.1 g	and/or	1180 s	1971 g	5.6 s	4.1 g
lamentation	<u>9</u>	<u>4</u>	7.8	5.9	heartache	<u>6</u>	<u>6</u>	6.4	6.0
anger	<u>55</u>	<u>151</u>	6.3	7.7	pain	<u>76</u>	<u>126</u>	5.4	6.1
joy	<u>106</u>	<u>27</u>	7.7	5.6	suffering	<u>22</u>	<u>20</u>	5.8	5.6

sadness	<u>27</u>	<u>41</u>	7.1	7.5	shame	<u>10</u>	<u>22</u>	4.7	5.7
guilt	<u>23</u>	<u>52</u>	6.1	7.2	disappointment	<u>6</u>	<u>18</u>	3.9	5.4
grief	<u>41</u>	<u>12</u>	7.1	5.2	confusion	<u>5</u>	<u>19</u>	3.5	5.4
regret	<u>28</u>	<u>8</u>	7.0	5.0	happiness	<u>14</u>	<u>4</u>	5.4	3.5
anguish	<u>5</u>	<u>19</u>	5.3	7.0	misery	<u>6</u>	<u>10</u>	4.6	5.2
rage	<u>18</u>	<u>44</u>	5.6	6.8	sympathy	<u>6</u>	<u>15</u>	4.0	5.2
remorse	<u>11</u>	<u>8</u>	6.8	6.0	distress	<u>7</u>	<u>10</u>	4.8	5.2
despair	<u>10</u>	<u>30</u>	5.3	6.8	loss	<u>27</u>	<u>115</u>	3.1	5.2
shock	<u>23</u>	<u>98</u>	4.6	6.7	frustration	<u>11</u>	<u>14</u>	4.8	5.1
loneliness	<u>6</u>	<u>15</u>	5.6	6.6					

From the table, it can be seen that there is a significant difference in salience (above 2) in *joy* (7.7/5.6), *shock* (4.6/6.7), and *loss* (3.1/5.2), which show although they occur in the *and/or* pattern of both *sorrow* and *grief*, *grief* is much more associated with *shock* and *loss*, and *sorrow* is much more significant in its co-occurrence with *joy*. The high frequency of co-occurrence of *shock* (98) with *grief* might imply that the trigger of the *grief* feeling may well be an unexpected event, leading to a sudden violent disturbance of the mind or emotions. We might assume that one's initial response to an immediate cause, a recent loss, often the death of a close person would be numbness and disbelief as one would feel it hard to accept the extremely unpleasant fact. From the statistics in the table, it can be argued that the principal determinant of *grief* is *loss* (115), which is well in line with psychological findings. The loss may be temporary (separation) or permanent (death), real or imagined, physical or psychological. The loss of a loved person often results in the experience of *grief*; losing a loved one also means the loss of strong bonds of friendship, companionship or love (Izard, 1991). From a psychological perspective, people who experience *grief* may come to blame themselves for the loss (*guilt*, 52); *anger* (151) and *despair* (30) may also occur as a result of blaming someone else for the separation or as a result of feeling being left alone or deserted (Izard, 1977).

Table 5.1.1 reveals that both *sorrow* and *grief* semantically prefer words denoting emotional responses, such as *anger/rage*, *pain*, *guilt/remorse/regret*, *anguish/distress*, *sadness*, *despair*, *loneliness*, *shame*, *disappointment/frustration* and *sympathy*. One reason for this would be that people's emotions are often complex. Take *grief* for example, the complexities of feelings that can be invoked in *grief*-contexts are clearly exhibited in the following examples from the BOE:

- 3) Instead, the families had to spend a long night of fear, worry, *grief* and total despair.
- 4) they arrange to leave Lahardane in the hands of Henry and Bridget, their housekeepers. But Lucy, with a passion for her home and the landscape around her, runs away. Torn by guilt and *grief*, her parents wait for news. When none comes they assume that they have driven Lucy to suicide and leave.
- 5) Then she broke down in tears again and buried her head in Maggie's shoulder. Banks took a deep breath. Claire's pain and guilt and *grief* were so real they broke over him in waves and made his breath catch in his chest.

It appears that the discussion of various feelings together in the discourse is quite common, as exemplified below:

- 6) It is not enough just to recognise what society has done to us leads us to feel shame, guilt, hate, *grief*, pride, envy and fear.

In other patterns of their sketch difference, significant differences in salience are also found. It shows that *private* (70/7, 4.8/1.5), *own* (/81/12, 3.4/0.6) and *share* (38/7, 4.6/2.2) occur much more frequently with *grief* than *sorrow*. It seems that *grief* is often a private matter, but it could also be a public and social one. *Sorrow* hardly co-occurs with *public* probably because it is more introverted and introversion implies being private. In the 'grief' only patterns, we find *shared* (7) in the 'a_modifier' pattern, *mother* (16), *family* (27) and *father* (6) in the 'possessor' pattern, and *overcome* (6) in

the 'object_of' pattern. This might suggest that *grief* is a feeling elicited by a misfortune which may be overcome by sharing it with others. For example, at a time of death, a family comes together to mourn, and although it is death that brings about this closeness, there is a sense of affectionate ties with other members of the family. Shared *grief* over the loss of a loved one can thus reunite a family and facilitate cohesiveness (Izard, 1991). *Grief* is a strong feeling which is usually expressed publicly. *Outpouring* (86, -5, +5⁹) of *grief* provides 'some relief and a sense of release' (Wellenkamp 1988: 495). Expressions of *grief* can also elicit empathy and strengthen bond among all those who are bereaved (Izard, 1977). *Mourning* (58, inclusive of *mourn*, -5, +5) is a typical context for *grief*. Its close association with *grief*, but not *sorrow*, indicates that *grief* is more a communal, ritualised, culturally-mediated and extroverted response to a bad event, in particular, someone's death.

I have searched for *grief is* in the BOE and got some of its paraphrases:

- 7) ...but the purpose of **grief** is to cope with the unacceptable, and drastic reactions are both necessary and healthy. If you care for someone in any way, you become attached to them. If you care for someone who dies, then you continue to love them - but there is no longer a person for you to love. We cannot alter our feelings as quickly as life and death can change the world. You may begin to feel foolish about your love for a 'no one', although it is unavoidable. The major grief feeling is LOSS. As Ann Carpenter and Geoffrey Johnson's book makes clear, we can sense this about absolutely anything that matters. We grieve in some way, often the same way, whether someone dies, or is taken from us by divorce. Or when we lose our pet cat, or move house, change schools, lose a favourite teacher or a 'best' possession. But of course the most severe loss occurs when we are bereaved by the death of someone close to us.

⁹ Word Sketch only presents salient patterns and will lead to some unavoidable omissions, so sometimes I also extend the span to -5, +5 and search for other meaningful collocates. In such cases, I will mark them after the frequency.

Naturally, the first grief feeling to hit when you get bad news is SHOCK and numbness.

- 8) When someone we love dies, **grief is** a natural consequence. It is a very strong emotion which can sweep us away so that our normal states of mind are eclipsed.
- 9) For most people, **grief is** not a chain-reaction in which we go smoothly from one state of mind to the next, but rather a switchback ride of feelings, often swinging rapidly from one state to another.
- 10) If **grief is** repressed or pushed down ('isn't she brave; she's behaving so normally ') it may emerge later as physical or depressive illness.
- 11) It is essential to be able to talk to each other and not bottle it in. Talking to other women who have miscarried is also very helpful, particularly if they have since succeeded in having a baby. This is also where Mum comes in. She can be your biggest comforter of all. If you do not talk to anyone or communicate, the **grief is** closely followed by anger, at yourself, at each other, and at the doctors who didn't explain; ` Surely they could have done something to stop me from miscarrying, although I know that it was...
- 12) At one level, **grief is** or can be shared - the level of support and understanding.
- 13) **grief is** a selfish emotion...
- 14) At a time when people feel isolated, with few common experiences to unite them, these expressions of grief can give them a sense of belonging, however momentary. So through these collective displays of emotion - showing the world how upset you are and connecting with others who feel the same way - an otherwise fragmented society is able to achieve a sense of collectivity . This can only ever be a temporary phenomenon - the **grief is** necessarily transient because, in a sense, it is not real. People may feel genuinely upset, but it is not authentic grief. Grief is the emotional expression of a profound sense of personal loss. So just as soon as today's public mourning springs up, it starts breaking up and disappears.

The above paraphrases give us almost the same information on *grief* as we get from its collocates. As *grief* is a much-talked about and well-researched topic in written discourse, we can get a good many paraphrases from the BOE which contains a multitude of books of life science. For other emotion expressions, it is quite not the case. But the case of *grief* shows that as long as we can collect sufficient paraphrases of a given lexical item, its meaning will become transparent.

So far, the discussion has focused on *grief* as a feeling. The following examples illustrate *grief* in its sense of ‘something that causes great unhappiness’, which only accounts for a tiny percentage of the total instances:

- 15) ...the love of money forces our thoughts in other directions. Consequently right in our own day, unless I am much mistaken, some of your books have disappeared, I fear beyond recovery. It is a great *grief* to me, a great disgrace to this generation, a great wrong done to posterity. The shame of failing to cultivate our own talents, thereby depriving the future of the fruits that they might have...
- 16) Although it had been known that Grant was in failing health, his death on May 13, 1902, in the midst of his work, at a crisis in Queen's affairs, came to the university not only as a *grief* and a loss, but as a grave blow.

Grief is conceived of as an emotion prototypically linked with death, so the death of someone who is special to a particular group might come as a *grief* as shown in Example 16. However, it can also be extended to other situations where one will suffer to a similar degree as consequence of other intolerably unpleasant events. The desertion of books and pursuit of money in Example 15 is a good illustrative example of such events.

The following examples illustrate what a *sorrow* in the sense of ‘something that causes great unhappiness’ could be (like *grief*, this sense of *sorrow* only accounts for a tiny percent of the total instances in the corpus):

- 17) Wertheimer: What does Magic Johnson's resignation mean to you, to the commission? What is the loss? <p> Rogers: Oh, it--it's the--it--of course, it's--it's a *sorrow* to all of us. It's a loss of a fine human being, a very articulate force.
- 18) You and I have been associated together for fourteen years - in name at least - and I had hoped for closer association in the future so that it is a very real *sorrow* for me to have the association end. And I can't help but ask why.

- 19) I noticed that she had aged more than the years warranted. She had lost her husband and then there had been the death of her daughter-in-law, Lucian's wife. That must have been a *sorrow* to her.
- 20) 'But her father died just before the wedding, which was a great *sorrow* to Ruth after the long separation from him.'
- 21) ...read to me and told me tales up to about my second year in elementary school, I think I would have been between four and seven. At that age, I had not yet known anything you could call a great *sorrow*. For that reason no doubt, when I learnt that in the end the little snail had stopped bemoaning his lot, I simply thought, 'oh, good # That was all. I gave no special thought to the whole matter...

Examples 17-21 show that a *sorrow* could be someone's resignation leading to 'the loss of a fine human being, a very articulate force', or the ending of a former association, or the loss of one's husband and daughter-in-law, or the death of one's father. It would seem that *sorrow* and *grief* talk about the emotional reactions towards someone's death at different stages.

As pointed out earlier, the emotional responses towards an unpleasant event or situation might be complex. The reactions may go through several phases: the first response would be the awareness of the loss; then one might think about the cause, the circumstance or the change it may bring about etc.; and finally the emotion would be transformed into a combination of *sorrow* with *rage*, *guilt*, *despair* or *anger*, or some (or all) of them, depending on the associations invoked by the given event. So it is hardly surprising to find that *sorrow* and *anger* often co-exist:

- 22) ...is that 300 people died in the bombing of the building where they took shelter, the building the US says was a military communications bunker. <p>Wertheimer: Reaction in the Arab world is both *sorrow* and anger. Jordan and Tunisia have declared days of mourning.
- 23) I must say that, like the solid citizen who took Julia Somerville's boyfriend's nudie snaps of her daughter to the police, I was all for contacting the boys in blue here. All I can say, more in *sorrow* than in anger, is what a strange picture of the British adult is emerging

here: the women using innocent babes for sensual gratification, and the men beating their tiny daughters on the bare...

24) Now, when she complains about the damp, the council says, 'You used to complain because you were four in a room in a mixed hostel, but now you can close the door behind you!' She says this more in *sorrow* than in anger. 'Even animals in the zoo, they treat them nicely. Who am I, what can I say, what can I do? Nothing. I am nothing, I count for nothing...

Examples 22-24 show that if you deal with something more in *sorrow* than in anger, you probably mean that you have considered the whole matter carefully and come to the conclusion that you are very sad or disappointed rather than angry about it. In Example 23, the speaker is extremely disappointed because he knows clearly what that terrible picture might suggest, and in Example 24, the speaker is very sad because she is aware how unimportant she is in other people's eye. Such conclusions require more thinking about what has happened. From the discussion above, it follows that *sorrow* involves a deeper thinking about everything around what has happened rather than a simple perceptual or emotional reaction. It is a more advanced and rational emotion involving 'higher' thought processes, as compared with *anger* or *grief*.

Sorrow is a 'feeling of sadness or distress caused esp. by loss, disappointment or regret' (OALECD). It implies that the experiencer considers the situation as uncontrollable and thus is not likely to attempt to do something about it. In other words, it is the awareness of powerlessness that is salient for *sorrow*. It suggests a degree of resignation ('I can't do anything about it'); it implies a 'semi-accepting', or 'semi-resigned' attitude towards what has happened (Wierzbicka 1999: 66). The pain comes not only from the irreparable loss, but also from what has been obtained from a rational analysis of the whole matter. This explains why people will be more in *sorrow* than in

anger when they clearly know that the outbursts of anger will not help. The most typical collocate of *sorrow* is *drown* (221), which can be explained by the idiomatic expression ‘drown one’s *sorrows*’. Let us look at more examples from the corpus:

- 25) An embassy employee who was present said: ‘We were slightly depressed when the first England goal went in, but then we decided we might as well drown our *sorrows* and forget it.’
- 26) The *Locust Room* is his first novel since *The Mercy Boys*, an exploration of maleness which had a cast of characters who drank in the same pub in a vain attempt to drown their *sorrows*.
- 27) Don't drink alone. If you use liquor at all, drink on social occasions only. If you use it to drown your *sorrows* and wash away your frustrations, liquor will use you--roughly.

As shown in the examples above, the phrase basically means one drinks a lot of alcohol in order to forget his/her troubles or problems, which could be some disappointment in Example 125, frustrations in Example 27 or something else. A *vain attempt* in Example 26 seems to show that it is not so easy, as they wish, to forget the problems by getting drunk. ‘In the case of *grief* and *grieving* the experience intentionally focuses on the painful subject (‘I want to think about this’), whereas in the case of *sorrow* there is rather, an inability to forget (‘I can’t not think about this’)’ (Wierzbicka, 1999: 67). *Sorrow* is a long-term suffering which cannot be forgotten easily because there is always a conflict between the consciousness of powerlessness and the reluctance to accept the unpleasant reality. *Grief* implies ‘poignant suffering for an immediate cause, a recent loss’ (Santangelo 2003: 389). ‘*Sorrow* may have its roots in the past, but the stress is on the on-going, long-term state’ (Wierzbicka 1999: 66). So it is likely that one still feels *sorrow*, which is less intense than *grief*, after even several years of the death of a close person. Example 21 seems to suggest that a young child does not know what

a *sorrow* is because s/he never gives a special thought to a sad matter due to the cognitive constraints.

Next let us look at *sorrowful* and *sorrowfully* in the corpus. The BOE contains 342 occurrences of *sorrowful*. By inspection of its word sketch, it is clear that *sorrowful* is often modified by *deeply* (7), *particularly* (3), and that it is semantically associated with things showing *sorrow*, such as *voice*(4), *eye* (5), *face* (4), *wail* (2), *look* (3), *sigh* (3), *glance* (2), *tear*(2), *sound* (2), *note*(2), or things causing *sorrow*, like *mystery* (6), *tale* (4), *sight* (3), *memory* (2). For *sorrowfully* (82), it seems that the pattern ‘shook (shaking) one’s head *sorrowfully*’ dominates the picture. So unlike *sorrow*, both *sorrowful* and *sorrowfully* are principally associated with external manifestations, namely more extroverted.

Now let us proceed to examine *grieve*. Its definitions are shown below:

LDCE:

①[I and T]to feel extremely sad, especially because someone you love has died

②[T] if something grieves you, it makes you feel very unhappy = upset

OALECD: *v* (*fml* 文)

①[Tn] cause great sorrow to (sb) 使 (某人) 极为悲伤: ②(a) [I, Ipr]

~ (for sb); ~ (over/about sb/sth) feel a deep sorrow because of loss

(因失去而) 感到悲痛: (b) [Ipr] ~ at/about/over sth feel deep regret (about sth) (对某事) 感到非常後悔、懊悔

COBUILD:

1 [VERB] V prep, V, V-ing

If you **grieve over** something, especially someone's death, you feel very sad about it.

2 [VERB] be V-ed by/at n, be V-ed to-inf, it V n to-inf, also V n

If you **are grieved by** something, it makes you unhappy or upset.

A simple glance of the definitions above shows that, basically, *grieve* means ‘feel *sorrow*’ and ‘cause *sorrow*’. Unlike the other two dictionaries, OALECD lists its sense

of 'feel deep regret (about sth)' as a separate one. Its collocate data (3675 hits) from the BOE show that its colligational items include prepositions, typically *for* (488), *to* (498), *over* (130), *at* (51), *by* (43), and modifiers, predominantly *deeply* (31), *properly* (23). Its top collocates also reveal that it is semantically associated with words indicating people involved in a grieving event (mostly someone one is very close to) such as *family* (235), *relative* (105), *parent* (85), *mother*(65), *widow*(60), *father*(38), *friend* (27), *husband* (23), *people* (29), *dad* (16), *woman* (18), *daughter* (12), *wife* (12), *brother* (11). Further, what can be seen from its top collocates is that it, in the form of *grieving*, prefers words related to a length or portion of time, like *process* (139) and *period* (18), which are all relevant to someone's death. From its word sketch, I find *loss* (20), *death* (7), *baby* (8), *son* (7) and *victim* (6) in the 'pp_for-p' pattern, and *privately* (6), *openly* (5), *alone* (10) and *together* (20) in the 'modifier' pattern. The former indicates that *grieving* usually happens due to someone's death, while the latter seems to show that *grieving* is an emotion characterised by both extroversion and introversion.

In sum, *grief* and *grieve* are more semantically associated with words indicating family members or someone with whom one has a very close relationship, words relevant to social rituals, public ceremonies, etc, while *sorrow* more semantically prefers a word describing whether the feeling is deep or sincere, such as *heartfelt*, *genuine*, *real*. The analysis seems to support Teubert's claim (2000) that *sorrow* is a feeling which is deeper, longer lasting, more introverted, more individual, more voluntary, more personal but less intense than *grief*; in contrast, *grief* is more extroverted, more communal, more ritualised and more culturally mediated than *sorrow*.

To find out whether the distinction between *sorrow* and *grief* is shared by Chinese, we need to examine their equivalents in Chinese texts.

5.2.1.2. Chinese equivalents of *sorrow** and *grief**

Queries of *sorrow** and *grief** are carried out in the BPC and the results are shown as follows:

Table 5.1. 2 Chinese equivalents for *sorrow** and *grief** in the BPC

	<i>sorrow</i> / 91 hits	<i>grief</i> / 72 hits
<i>beishang</i>	27	23
<i>you/chou</i> #	9	---
<i>shangxin</i>	9	10
<i>nanguo</i>	6	---
<i>beiai</i>	---	8
<i>beitong</i>	---	7

--- Less frequent in the parallel corpus

Indicating a group of words denoting the *you/chou* (worry) concept

Their Chinese equivalents include *beishang*, *shangxin*, *nanguo*, *beiai* and *beitong*, excluding *you/chou*. In fact, *shangxin* and *nanguo* describe less intense sad feelings. They were used as translations of *sorrow* or *grief* due to the simplification in translating (cf. Baker, 1995). I will discuss them in the next section. The concept of *you/chou* focuses on worry, which is not my focus, so in this section, I will only discuss *beishang*, *beitong* and *beiai*. Before investigating further, let us look at their definitions in XHCBL, which are shown below:

Table 5.1. 3 Definitions of *beishang*, *beitong* and *beiai* in XHCBL

	Chinese definition	English definition
<i>beishang</i>	<i>shangxin</i> ; <i>nanguo</i>	<i>sad</i> ; <i>sorrowful</i> ; <i>mournful</i>
<i>beitong</i>	<i>shangxin</i>	<i>sorrowful</i> ; <i>grieved</i>
<i>beiai</i>	<i>shangxin</i>	<i>sad</i> ; <i>sorrow</i> ; <i>grief</i>

The table seems to suggest that *shangxin* is a generic word to denote a sad feeling, but it tells us nothing about their distinctions. Before analysing the three terms starting with

bei -- *beitong*, *beiai*, and *beishang*, let us take a look at the Chinese system of lexicon. Modern Chinese words (polysyllabic *ci*) evolved from old Chinese characters (monosyllabic *zi*). In old Chinese, each single character (*zi*) has its own self-contained semantic meaning, such as *bei*, *ai*, *tong*, *shang* etc. The vast majority of modern Chinese words come from the combination of two or more old characters and have developed their new denotations and connotations over time, but there are still close links between the old meanings of the constituents and the new meaning of the modern word, especially in some conventionalised phrases and idioms. Yip (2000) gives a detailed account of the structure of Chinese lexicon:

Chinese is not a language totally deprived of morphological derivations, it is perhaps far less morphologically prone than a language such as English. Generally speaking, instead of being composed of 'base+affix', a Chinese di- (or poly-) syllabic lexeme will, in the majority of cases, assume the look of a compound, that is, with the two (or more) constituent mononyms contracting a kind of quasi-syntactic relationship with each other, which makes it possible to analyse the internal composition of a Chinese word in syntactic terms (p. 90).

One of the syntactic structures found in disyllabic lexemes is juxtapositional type (juxtaposer+juxtaposed/juxtaposed+juxtaposer), in which the combination of two mononyms of similar orientation and syntactic category forms a word. The primary motivation behind adopting a juxtapositional structure is perhaps to conform to the overall disyllabification tendency of the modern lexicon. In such a set,

the distinctive feature of a juxtaposer is actually its juxtapositional capacity. It is the mononym that finds a wider distribution and occurs more frequently in a group of similar juxtapositions. This is particularly true for adjectives. In a semantically related set of disyllabic adjectives, the juxtaposer is usually the one which occurs more frequently or enjoys greater combinatory power and therefore defines and charts the general semantic orientation of the set, whereas the juxtaposed is the one which tends to make more minute differentiations of meaning within the same

general semantic orientation set down by the juxtaposer or determine the syntagmatic potentiality of the resultant combination (Yip, 2000: 95).

In the case of the *bei* set, obviously, *bei* is the juxtaposer, which sets the tone for the whole set, and the other three mononyms are the juxtaposed, fine-tuning the meaning of the juxtaposer. I will try to tease out the semantic differences between this lexical set. As noted earlier, Chinese feeling-denoting words are usually seen and glossed as adjectives, but they can also be used as nouns and adverbs. The present Chinese POS tagging software cannot annotate their word classes with sufficient accuracy, so I chose not to tag them at all. My analysis will focus on the semantic content, i.e. the concept, so I will discuss them together.

5.2.1.2.1 *Beitong*

The significant collocates of *beitong* (2677¹⁰ hits) are shown below:

Table 5.1. 4 Significant collocates of *beitong* in the CCL corpus

Introversion/ extroversion	Cause	Modifier	People involved	Others
<i>gandao</i> 121	<i>si</i> 75	<i>wanfen</i> 117	<i>muqin</i> 55	<i>lingren</i> 20
<i>shen'gan</i> 55	<i>shishi</i> 54	<i>juda</i> 99	<i>fuqin</i> 27	<i>ren</i> 92 (<i>qiangren</i> 52)
<i>xinqing</i> 79	<i>qushi</i> 41	<i>jidu</i> 42	<i>fumu</i> 23	<i>chenjin</i> 69
<i>xinli/xinzhong</i> 36/33	<i>bingshi</i> 19	<i>shifen</i> 74	<i>qinren</i> 21	<i>xianru</i> 24
<i>neixin</i> 31 <i>xin</i> 28	<i>xisheng</i> 21	<i>hen</i> 69	<i>qinshu</i> 22	<i>chongman</i> 19
<i>biaoshi</i> 55	<i>yu'nan</i> 42	<i>feichang</i> 63	<i>jiashu</i> 17	<i>zhenjing</i> 39
<i>shenbiao</i> 21	<i>e'hao</i> 32	<i>shenshen</i> 22	<i>zhangfu</i> 22	<i>xiaoxi</i> 42
<i>ku/kuqi</i> 46/19	<i>shiqu</i> 57	<i>wuxian</i> 30	<i>qizi</i> 18	<i>fennu</i> 37
<i>shengyin</i> 19	<i>buxing</i> 38	<i>buyi</i> 29	<i>erzi</i> 20	<i>juewang</i> 24
<i>leishui/yanlei</i> 24/19		<i>wubi</i> 17		

From the table, we can find that *shen'gan* (deeply+feel) and *shenbiao* (deeply+ express) are segmented as one word by ICTCLAS. They occur very frequently, but they are only recurrent patterns rather than single words. We will ignore this inaccuracy of word

¹⁰ The frequency generated by Wordsmith is not necessarily the same with the one by the CCL tool because the latter counts the number of lines containing the node word, regardless of how many times it occurs in the same line, while the former takes the real number of occurrences of a word as its frequency. For example, 伤感 in the line ‘我要对你说的是底下的话 这些话也许能安慰你,也许更使你[伤感], 由[伤感]而得病’ is counted as 2 occurrences by Wordsmith, but only 1 by the CCL tool.

segmentation, and focus our analysis on meaning. The first column lists the words pointing to its introversion, such as *gandao/(shen)gan* (feel), *xinqing* (frame of mind), *xinli/xinzhong* (in the heart), *neixin/xin* (inner heart/heart), and words showing its feature of extroversion, like *biaoshi/(shen)biao* (express), *ku/kuqi* (cry/weep), *shengyin* (voice), *yanlei/leishui* (tears). English emotion words also frequently collocate with *feel*, but hardly with *in the (inner) heart*; Chinese tends to use *xinli/xinzhong* or *neixin* to emphasise that what is talked about is an inner feeling. I will address the cultural reason for this striking difference in detail in Section 5.2.2. Column 2 shows the causes for *beitong*: *si/shishi/qushi* (die/death), *bingshi* (die of disease), *xisheng* (give one's life for/die a martyr's death), *yu'nan* (be killed in a disaster), *e'hao* (news about the death of a beloved person), *shiqu* (loss/lose), and *buxing* (misfortune). *Buxing* is typically related to death when it is associated with *beitong*. Hence, it may be said that the predominant trigger for *beitong* is someone's death. Since Chinese emotion words can be used as a noun, an adjective or adverb, we find two classes of modifiers in the third column, i.e. adjectival, like *juda* (enormous), and adverbial, such as *wanfen/ji(du)/wubi* (extremely), *shifen/hen/feichang/* (very), *shen(shen)* (deeply), *wuxian/buyi* (endlessly). According to the fourth column, the people involved in the feeling of *beitong* include *muqin* (mother), *fuqin* (father), *fumu* (parents), *qinren/qinshu/jiashu* (family), *qizi* (wife), *zhangfu* (husband), and *erzi* (son), who are either the object or the experiencer of *beitong*. The last column lists its other collocates which might contribute to the unravelling of its meaning. In Chinese, in most cases, emotion words have to be preceded by 'lingren/jiaoren/rangren/shiren' (make/cause) to express a causative meaning. For

instance, if you want to talk about *sad news* which makes people *beitong*, you have to say ‘lingren *beitong* de xiaoxi (news)’. For the same complex concept, English uses a single word to express it, but Chinese uses a phrase instead. The former is referred to as ‘synthetic expression’ and the latter ‘analytic expression’ (Banczerowski, 1980: 336). We will return to this issue later. *Ren* (control) might suggest people sometimes control their feelings and prevent them from exploding. 52 occurrences of *qiangren* (be forced to control) reveal that, in most cases, people are forced to do so. *Beitong* also collocates with strong negative feelings, such as *fennu* (indignation), *juewang* (despair). *Chenjin* (be immersed in) and *chongman* (be filled with) might show that *beitong*, in Chinese, is metaphorically conceptualised in terms of fluid. The association of a strong emotion with the state of ‘be immersed in’ is specific to Chinese. We will come back to this issue later.

From its collocates *zhenjing* (shock), *xiaoxi* (news), and *e’hao* (news about the death of a beloved person), we can see that *beitong* is more an overwhelming emotion erupting like a volcano, invoked by something extremely bad and unexpected you are unwilling to face or find hard to accept, prototypically by the death of a close person. It is an intense feeling and usually accompanied by some physiological reactions, such as crying/weeping and tears. 262 hits of *beitong yujue*, a 4-character idiom, are found in the CCL corpus:

beitong yu jue
beitong on the point of stopping breathing
‘so *beitong* as to come to the end of life’¹¹

¹¹ The English glosses in single quotation marks are provided by the author, henceforth.

From this idiom, it can be inferred how strong the feeling of *beitong* is. *Beitong* may come suddenly and fade away over time. It comes from the bottom of the heart but is often shown outwardly. So we can argue that *beitong* is more individual than communal, more voluntary than ritualised, more personal than culturally mediated. It is similar to *sorrow* in terms of these three dimensions, but to *grief* in terms of extroversion, intensity and length.

5.2.1.2.2 *Beiai*

The top collocates of *beiai* (3281 hits) are shown as follows:

Table 5.1. 5 Significant collocates of *beiai* in the CCL corpus

Introversion	Extroversion	Modifier	Feeling	Others
<i>gan(dao)</i> 248	<i>kuqi</i> 26	<i>feichang</i> 33	<i>tongku</i> 100	<i>ling</i> 46
<i>juede</i> 51	<i>ku</i> 39	<i>zhen</i> 33	<i>fennu</i> 58	<i>si</i> 43
<i>neixin</i> 21	<i>yanjing</i> 28	<i>shifen</i> 25	<i>shiwang</i> 33	<i>shiqu</i> 27
<i>xinli</i> 49	<i>yanlei</i> 25	<i>shen(shen)</i> 42	<i>juewang</i> 30	<i>buxing</i> 19
<i>xinzhong</i> 36	<i>lian(shang)</i> 24	<i>juda</i> 20	<i>jimo</i> 27	<i>chongman</i> 43
	<i>shengyin</i> 37	<i>moda</i> 20	<i>gudu</i> 21	<i>yizhen</i> 29
	<i>biaoqing</i> 18	<i>wuxian</i> 18	<i>kongju</i> 26	<i>minzu</i> 29
	<i>shenqing</i> 18	<i>yidian</i> 29	<i>huanle</i> 42	
		<i>xie/si</i> 25/20	<i>kuaile</i> 24	

From the table, it can be seen that *beiai* is not only a mental state of emotion, as indicated by *gan(dao)/juede* (feel), *neixin* (inner heart) and *xinli/xinzhong* (in the heart), but is also characterised by outward expression, as revealed in *kuqi/ku* (cry/weep), *yanjing* (eyes), *yanlei* (tears), *shengyin* (voice), *biaoqing/shenqing* (facial expression) and *lian(shang)* (on the face). Its frequent modifiers are *feichang/shifen* (very), *zhen* (really), *shen(shen)* (deeply), *juda* (enormous), *moda* (utmost), and *yidian/xie/si* (a little). The feelings in its company include *tongku* (anguish), *fennu* (indignation), *shiwang* (disappointment), *juewang* (despair), *kongju* (fear), *jimo/gudu* (loneliness), as well as some positive feelings, like *huanle* (joy) and *kuaile* (gaiety). *Ling* (cause) occurs in

beiai's collocates for the same reason for *beitong*. *Chongman* (be filled with) *beiai* is also very common. *Yizhen* (lasting for a short time) seems to show that *beiai* may come suddenly but strongly and disappear after a very short time. *Si* (death), *buxing* (misfortune) and *shiqu* (loss) may cause *beiai*, but unlike *beitong*, they are not predominant triggers. In addition, I found quite a few hits of *minzu* (nation) *de beiai*, *rensheng* (life) *de beiai*, *renlei* (human kind) *de beiai*. The concordance listing for *minzu de beiai* is generated from the CCL corpus:

Figure 5.1 Concordance listing of *minzu de beiai*

N	Concordance	et	agl	ord	#	s	%
7	bt 文章标题: 作者: 1562: 我常常 [悲哀] 地感到, 我们的民族过于健忘。"文革" 不过十年, 已经很少再见提				1,450	871	%
8	师大家。像陈寅恪、梁思成那样, 他们后来的遭际, 是民族的厄运和 [悲哀]。这后一半, 它出了许多大官。像				7,988	1911	%
9	、横厉, 而又 [悲哀] 至于毁灭点的节奏——我及我的民族是受到了极度的, 人世再无以复加的创伤, 且无				5,881	1101	%
10	月份 a. bt 文章标题: 作者: 376: 日救亡教育, 对当前的民族灾难和应有的态度有较好的认识。他们不 [悲哀],				9,759	1919	%
11	标题: 作者: 814: 判、即否定自身的能力所致。一个民族、一个国家、一个社会最大的 [悲哀] 之一就是自我				7,482	1911	%
12	中国革命战士的精神风貌, 因而这部书也反映出了一个民族的大 [悲哀] 与大欢乐, 具有厚重的内涵和史料价值				8,289	1919	%
13	却最终落后挨打, 这是中国发明史的 [悲哀], 也是一个民族因封建制度造成的停滞带来的无奈。【文件名: \当代				6,347	1919	%
14	代价, 如果高尚必须迁就和宽容卑鄙, 那无疑是一个民族的巨大 [悲哀]。【文件名: \当代\报刊\读者\读者 (6,453	1101	%
15	: 文学长短篇小说作者: 曾卓 1727: 不奇怪: 它们把西班牙民族的心理概括住了; 它们表明了听天由命, [悲哀],				0,579	1251	%
16	13 月份 a. bt 文章标题: 作者: 344: 咫尺千里——民族的悲剧, 国人的 [悲哀]。【文件名: \当代\报刊\				7,733	1919	%
17	涌来"心甘情愿"地接受"心理征服", 这不能不说是一种民族的 [悲哀]。无怪乎日本方面对中国留学生数目激增持				4,033	1921	%
18	未来, 文化如果沦落到"让人玩儿"的地步, 将是一个民族的 [悲哀], 他希望更多的人能尽快认识到这一点。				5,457	1521	%
19	我希望第5、6、7届不要再办下去了, 因为这是我们民族的 [悲哀], 为什么会有海峡情呢? 就是一道海峡把				6,880	1731	%
20	民族只会看小人书, 而看不了《红楼梦》, 是这个民族的 [悲哀]。交响乐永远不可能流行, 如果听交响音乐				4,060	1919	%
21	: 作者: 853: 败坏于他人, 或许是发明了围棋的那个民族的 [悲哀]; 但找到了第二故乡并因此而繁荣发达,				9,854	1911	%
22	对什么都无所谓的态度, 不仅是个人的 [悲哀], 也是民族的 [悲哀]。【文件名: \当代\报刊\作家文摘				8,038	1919	%
23	的出现上。这是唯心史观在爱国主义观上的反映, 也是民族的 [悲哀]。一些民族英雄、爱国志士之所以"空怀移				9,164	1919	%
24	是中国知识界个人的 [悲哀], 更确切地说, 倒是整个民族的 [悲哀]。中国文化之所以在近代不能对【文件名: \				6,426	1911	%
25	1658: 果真有这么一天, 那就不仅是他们的不幸, 更是全民族的 [悲哀]了。【文件名: \当代\文学\大陆作家\张正				6,218	1981	%
26	: 作者: 216: 余秋雨说: "失落了上海的文明, 是全民族的 [悲哀]。【文件名: \当代\报刊\1994年报刊精选				1,070	1919	%
27	不积极努力避免, 那就不仅仅是文人的 [悲哀], 更是全民族的 [悲哀]了。【文件名: \当代\文学\大陆作家\曾卓				0,816	1251	%
28	希望第5、6、7届不要再办下去了, 因为这是我们民族的 [悲哀], 为什么会有海峡情呢? 就是一道海峡把				8,752	1919	%
29	: 作者: 153: 从这个意义上说, 科教片受冷遇是我们民族的 [悲哀], 而振兴科教片则关系振兴中华、振兴民族				7,992	1919	%
30	: 123: 这些不正常的现象形成习惯之后, 那将是一个民族的 [悲哀]。【文件名: \当代\报刊\1994年报刊精选\01				6,578	1919	%
31	未来, 文化如果沦落到"让人玩儿"的地步, 将是一个民族的 [悲哀]。【文件名: \当代\报刊\1994年报刊精选\05				8,050	1919	%
32	没有的。原因是: 他关心祖国的前途, 他尊重自己的民族文化。【文件名: \当代\报刊\读书\wol-139. txt 文章标题				1,802	1911	%

From the concordance listing, it can be seen that it is seen as *minzu de beiai* if its culture has been treated lightly, if its people only read picture-storybooks and cannot understand *Honglouloumeng*, if indifference has become prevalent in that nation, if scientific and educational films are given the cold shoulder, if abnormal immoral phenomena have become normal, etc. It seems that, in modern texts, *beiai*'s original

meaning of denoting the sad emotion associated with death or calamity has given way to its new sense to describe something which is sad, disappointing but we can do nothing about, something which should not have happened, something which will lead to very bad consequences if it continues to be ignored, especially when people are not aware of its seriousness at all. The concept of *beiai* has gone beyond its original circle and been extended to situations connected with a wider range of unfortunate events happening to people, nations, life, human beings or even the whole world. A young girl who is not good-looking may be shrouded in *beiai* if she finds no boy notices her existence at all. If we feel *beiai* for ourselves, we are fully aware what we did or what has happened to us is really not something worth being proud of or something which should be encouraged, but something which will turn out to be the opposite. It is a *beiai* thing if a young girl abandons herself to despair, if a nation does not give some serious thought to its history, if people think love is worth nothing compared with money, if a person even cheats his or her own mother, if the movie industry is only seeking commercial profits without thinking about the negative influence it may have on the society, or if a writer produces rubbish writings just for money or political benefits etc. From the above analysis, we may argue that *beiai* seems to be caused by regret, sympathy, pity or disappointment. Consider the following examples:

- 28) 这已不是讽刺,已经是悲哀,一种人们只有在接近绝望时才会感到的悲哀(*beiai*)。‘This is already *beiai*, not irony any more, which people can feel only when they are close to despair.’¹²
- 29) 无可奈何的悲哀(*beiai*) ‘*beiai* due to one’s powerlessness or helplessness’

¹² Throughout the thesis, if the English part of a parallel example is put in single quotation marks, it means that, by default, the example is from the CCL corpus and the translation is provided by the author.

- 30) 无心或心灰意懒是人生最大的**悲哀**(*beiai*)，因而庄子说过，哀莫大于心死。‘The greatest *beiai* for one’s life is that one feels disheartened and has lost the interest to do anything. Just as Zhuangzi said, the utmost *ai* is *xinsi* (lit. dead heart/ broken spirit).’

Disappointment or even despair is more clearly revealed in the above examples.

In its original sense, *beiai* still keeps its basic flavour of the deep sad stir, though not exclusively, due to the loss of a loved person. *Beiai* is more altruistic than personal, for its focus is on other people, not on the experiencer. It tells us more about what the experiencer thinks about a particular sad event. It implies ‘sympathy, compassion or even love, revealing a wish for bad things not to happen to the other person and a desire to do something good for the unfortunate’ (Ye, 2001: 375). This also explains how its modern sense has developed, for its new sense is also altruistic, compassionate, sympathetic, focusing on other people or events and revealing a wish for unpleasant things not to have happened, as noted earlier. The following examples bear out evidently this altruistic, compassionate, communal, ritualised and more extroverted emotion concept:

- 31) 在餐室里，她那么平静而又拘谨，而现在，她脸上一切伪装的**悲哀**(*beiai*)都已烟消云散，双眼闪烁着生活欢乐的光辉，面部被同伴的妙语逗乐的笑纹未消。In the dining-room she had been demure and discreet. Now all pretence of *grief* had passed away from her. Her eyes shone with the joy of living, and her face still quivered with amusement at some remark of her companion.¹³
- 32) 许多东西会在她心里引起**悲哀**(*beiai*)—那些弱者,那些凄苦无依的人，一概激起她的伤心。Sorrow in her was aroused by many a spectacle—an uncritical upwelling of *grief* for the weak and the helpless.
- 33) 趣味要高尚一点,不要用那条纱巾来表现自己实际上从来没有过的**悲哀**(*beiai*)。And better taste than to wear that veil to advertise a *grief* I’m sure you never felt.

¹³ If the English part of a parallel example is not put in single quotation marks, it means it is from the BPC, unless otherwise stated, henceforth.

- 34) 她一定要把自己渴望的勇气鼓起来，到牧师住宅去打听消息，对他的沉默表示自己的悲哀(*beiai*). Surely she might summon the courage of solicitude, call at the Vicarage for intelligence, and express her *grief* at his silence.

All *pretence of* in (31), *advertise* in (33), and *express* in (34) allude to the altruistic, extroverted and communal aspects of *beiai*, and *that veil* in (33), probably to ritualised aspect. In (32), *the weak and the helpless* points to its altruistic, compassionate aspects. And Example 32 seems to suggest that *beiai* could be a voluntary feeling elicited either in death-related or other situations.

5.2.1.2.3 *Beishang*

Beishang shares the common semantic core with both *sorrow* and *grief* since 27 out of 91 occurrences of *sorrow* and 23 out of 72 occurrences of *grief* have been translated as *beishang*. 2603 hits of *beishang* are found in the CCL corpus. Its significant collocates are given below:

Table 5.1. 6 Significant collocates of *beishang*

<i>Introversion</i>	<i>Extroversion</i>	<i>Modifier & quantifier</i>	<i>Feeling</i>	<i>Others</i>
<i>gandao</i> 102	<i>lian</i> 20	<i>hen</i> 91	<i>tongku</i> 117	<i>ling</i> 34
<i>juede</i> 30	<i>lianshang</i> 25	<i>feichang</i> 30	<i>fennu</i> 54	<i>rang</i> 29
<i>ganjue</i> 17	<i>yanjing</i> 38	<i>shifen</i> 29	<i>juewang</i> 29	<i>chongman</i> 55
<i>xinzhong</i> 35	<i>yanlei</i> 35	<i>jidu</i> 27	<i>kongju</i> 21	<i>si</i> 50
<i>xinli</i> 30	<i>lei</i> 18	<i>buyi</i> 26	<i>fannao</i> 19	<i>tanxi</i> 21
<i>neixin</i> 20	<i>shenqing</i> 25	<i>wuxian</i> 21	<i>youyu</i> 19	<i>muqin</i> 38
<i>xin</i> 39	<i>biaoqing</i> 22	<i>guodu</i> 38	<i>youchou</i> 18	<i>fuqin</i> 18
<i>qingxu</i> 28	<i>yangzi</i> 21	<i>yidian</i> 18	<i>shiwang</i> 18	
<i>xinqing</i> 18	<i>xiande</i> 20		<i>xiwqng</i> 20	
<i>ganqing</i> 20	<i>ku</i> 47		<i>xiyue</i> 17	
	<i>kuqi</i> 25		<i>huanle</i> 29	
	<i>shengyin</i> 21		<i>kuaille</i> 32	
			<i>gaoxing</i> 17	

A simple glance of the table shows that *beishang* shares most of the collocates with *beiai*. Like *beitong* and *beiai*, *beishang* is also conceptualised as fluid or gas because it frequently collocates with *chongman* (be filled with). The main differences lie in *beishang*'s unique collocations with *guodu* (excessively), as found in the pattern

'*beishang guodu/guodu beishang*', and with *fannaο* (vexation), *youyu* (melancholy), *youchou* (worries) and *tanxi* (sigh). *Tanxi* 'conveys, to a certain extent, the powerlessness and resignation on the part of the experiencer' (Ye, 2001, 366). *Beishang* implies that the experiencer has no other choice but to accept what has happened, however bad it is. It is an emotion not as violent as *beitong*, if it is elicited by the death of a close relative or a friend or even a national leader. However, it is more associated with something that is lost but you are longing to have it again. It is more related to *muqin* (mother) probably because Chinese mothers tend to be *beishang* when they feel powerless. One may feel *beishang* when s/he realises no progress has been made in his or her career after s/he struggle for ten years or when he feels inferior to others after losing the wife, family and job in his forties. Even a young sentimental boy will feel *beishang* if he has lost his favourite toy. Like *beiai*, '*wukennaihe* (powerless/helpless) de *beishang*' (7 hits) was also found in the corpus. *Beishang* is also a feeling triggered by parting, which is something people do not want to face but have no choice.

Beishang is a sentimental concept which highlights the sense of powerlessness due to the laws of nature and society. It is more an emotional state closely related to the attitude toward life and the world, more tragic and fatalistic, more a result of accumulation after experiencing a series of unfortunate events. It is something heavy that settles at the bottom of the heart after experiencing misfortunes. The following paraphrase from the CCL seems to support the claim that *beishang* is a long lingering haunting thought left after a sad experience:

- 35) 悲伤(*beishang*)是不能够匆匆而过的...
'Beishang is not a fleeting experience'

Beishang usually arises from the bottom of the heart and can be buried and hidden deep in the heart for the time being but stirred again by any possible trigger. It involves a reflection quality after the stirring. The following examples from the parallel corpus may shed light on the haunting and sentimental characteristics of *beishang*:

- 36) 同时他也按照他混乱的和激动的思想推理，认为苔丝只是在她提到的过度**悲伤**(*beishang*)下一时失去了心理平衡，才陷入这种深渊的。As well as his confused and excited ideas could reason, he supposed that in the moment of mad *grief* of which she spoke her mind had lost its balance, and plunged her into this abyss.
- 37) 她能向瑞德说些什么才可以去缓解他的**悲伤**(*beishang*)和恢复他的理智呢？
What could she say to Rhett that would ease his *grief* and bring him back to reasoning?
- 38) 苏伦和卡琳哭了一阵睡着了，她们每天至少要来这么两次，因为一想起母亲便感到**悲伤**(*beishang*)，觉得自己孤苦无依，眼泪使簌簌地从深陷的两腮上往下流。Suellen and Careen had cried themselves to sleep, as they did at least twice a day when they thought of Ellen, tears of *grief* and weakness oozing down their sunken cheek.
- 39) 他向她走来时，思嘉发现他眼中充满了**悲伤**(*beishang*)，同时也含有厌恶和轻蔑之情，这使她惊慌的心里顿时涌起满怀内疚。As he came towards her, she saw that there was *grief* in his eyes and also dislike and contempt that flooded her frightened heart with guilt.
- 40) 这里有过她的罪孽，这里有过她的**悲伤**(*beishang*)，这里也还会有她的忏悔。Here had been her sin; here, her **sorrow**; and here was yet to be her penitence
- 41) 多少希望和惋惜，多少**悲伤**(*beishang*)和痛苦。
Oh, ye legions of hope and pity—of **sorrow** and pain
- 42) 经理觉得这哀诉是对他个人而发，就好像他们俩单独在一起，他几乎忍不住要为他所爱的女子流泪，她是那么孤弱无助，那么**悲伤**(*beishang*)凄婉，又那么妩媚动人，楚楚可怜。The manager suffered this as a personal appeal. It came to him as if they were alone, and he could hardly restrain the tears for **sorrow** over the hopeless, pathetic, and yet dainty and appealing woman whom he loved.
- 43) 她也想到赫斯渥，但是想到他，只给她带来**悲伤**(*beishang*)。
Nothing but **sorrow** was brought her by thoughts of Hurstwood
- 44) 除了她自己而外，谁也没有看出故事中的**悲伤**(*beishang*)来。
None of them but herself seemed to see the **sorrow** of it

45) 一个人在**悲伤**(*beishang*)停止的时候，睡眠就会乘虚而入。

When *sorrow* ceases to be speculative sleep sees her opportunity.

Grief seems to be more closely associated with someone's death. In (38), the thought of their late mother made them feel *grief*, while in (43), the thought of Hurstwood would only made Sister Carrie feel *sorrow*. Both of them have been translated as *beishang*, but this *beishang* is not that *beishang*, which are not distinguished in the Chinese lexicon. The former is a more intense feeling, a fading version of *beitong* over time, emphasizing more on the missing element, and the latter is more lingering, haunting, sentimental and fatalistic, which explains why in (40), (41), (44), (45), *sorrow*, a more general version of *beishang* in Chinese sense, is used instead of *grief*. The *mad grief* in (36) is inferior in intensity to the *beitong* of losing of a loved one recently. In (37) and (39), *ease his grief* and *grief in his eyes* seem to suggest that *beishang* is something haunting and introverted you need to read from the face including eyes. Rhett's explosive *beitong* towards his daughter's death has developed into longer lasting *beishang*. It takes much longer time to overcome *beishang* than *beitong*. In (42), *the tears for sorrow*, which is quite different from *the tears of grief* in (38), reflects a sentimental thought typically related to powerlessness, helplessness, parting or loneliness.

5.2.1.3 Comparison

In this section, an in-depth contrastive-semantic analysis of *sorrow* & *grief* as well as their Chinese equivalents—*beishang*, *beiai* and *beitong* (excluding the *you/chou* concept) has been made. Drawing on the statistical data and typical parallel examples from the corpora, the discussion focuses on the distinctions between these emotions in

cause, intensity, length and accompanying behaviour etc. *Beitong* is an explosive intense sadness typically associated with someone's death. *Beiai* is an altruistic, sympathetic and compassionate feeling focusing on other people or events. *Beishang* could be a less intense version than *beitong*, which will develop into *beishang* over time, or a long lingering haunting thought as a result of accumulation of unfortunate experiences, closely related to a fatalistic outlook. The lexical set of sadness shares basically the same semantic core but each focuses on different elements. Table 5.1.7 summarises their features based on the previous discussion:

Table 5.1.7 Contrastive analysis: *sorrow & grief* vs *beitong, beishang & beiai*

	Extroverted	Communal	Altruistic	Explosive	Intense
<i>beishang</i>					less
<i>beiai</i>	√	√	√		less
<i>beitong</i>	√			√	√
<i>sorrow</i>					less
<i>grief</i>	√	√		√	√

Obviously, the concepts of emotions are categorised differently in different languages.

The Chinese language distinguishes sadness between whether it is extroverted or introverted, whether it is individual or communal, whether it is altruistic or personal, whether it is explosive or long lingering, or whether it is intense or less, while English lacks the dimension of being altruistic or personal. The feature of being altruistic is Chinese-specific in the domain of emotion. As discussed earlier, three different Chinese sadness concepts (*beishang*, *beiai*, and *beitong*) are encoded in the same word *grief* in English. This is because *grief* overlaps much with *beitong* for their both being extroverted, personal, explosive and intense; *beiai* was selected by the translators to express the concept of *grief* because both are extroverted and communal, though *beiai* is

less intense and altruistic. From the above table, it follows that there is no exact Chinese equivalent to *grief*. *Beiai* is the only choice when the element of being communal is emphasised. As shown earlier, *beishang*, a less intense feeling than *beitong*, more introverted, more personal, more individual, longer lasting feeling than *beiai*, is shared by *sorrow* and *grief*. *Beishang* is similar to *sorrow* in terms of being more introverted, individual, voluntary, lingering, haunting and less intense. However, as noted earlier, it also denotes a rather intense feeling, which is a fading version of *beitong* over time, emphasizing more on the missing element. This is why it was chosen to translate *grief*.

5.2.2 *Unhappy** and their Chinese equivalents

5.2.2.1 *Unhappy**

Before discussing *unhappy*, we need to look at what is *happy*. COBUILD says that ‘someone who is happy has feelings of pleasure, usually because something nice has happened or because they feel satisfied with their life’, and that ‘a happy time, place, or relationship is full of happy feelings and pleasant experiences, or has an atmosphere in which people feel happy’. So *happy* in English is used to talk about someone’s pleasant feeling or something which can lead people to such a feeling. Now we have a basic idea about what *happy* is, but an antonym with a negative prefix is not necessarily the exactly opposite to its stem in meaning. Let us look at how *unhappy* is defined in the three dictionaries:

LDCE:

☐ not happy ☐ feeling worried or annoyed because you do not like what is happening in a particular situation ☐ *formal* an unhappy remark, situation etc is not suitable, lucky, or desirable

OALECD:

☐ (a) sad or miserable; not happy 悲伤的; 难过的; 不幸福的; 不愉快的 (b) ~ (about/at sth) anxious or dissatisfied 忧虑的; 发愁的; 不满意

的 ▢unfortunate or unlucky; regrettable 不幸的; 不走运的; 令人遗憾的
的 ▢[usu attrib] (*fml*) not suitable or appropriate 不合适的; 不恰当的
COBUILD:

1 [ADJ] If you are **unhappy**, you are sad and depressed.

2 [ADJ] v-link ADJ, oft ADJ about/at n/-ing, ADJ that

If you are **unhappy about** something, you are not pleased about it or not satisfied with it.

3 [ADJ] ADJ n An **unhappy** situation or choice is not satisfactory or desirable.

A simple glance of the senses of *unhappy* reveals that an *unhappy* person feels no pleasure and an *unhappy* thing gives no pleasure. Its meaning of ‘feeling worried or annoyed’ is reflected in the first two dictionaries but defined in different ways. COBUILD does not list its sense of ‘unfortunate or unlucky’ as a separate one probably because this sense is not frequent in contemporary discourse. One can be *unhappy* about many things, but in real language use it has its own preferences to modify. To find out which words it prefers, we need to look at its corpus data.

Unhappy occurs 9611 times in the BOE and its top 50 collocates tell us that *unhappy* often precedes *marriage* (146), *people* (121), *childhood* (75), *memory* (65), *experience* (45), *man* (57), *family* (38), *customer* (37), *life* (37), *relationship* (33), *player* (33) and *ending* (31). *Marriage* is the most frequent noun modified by *unhappy* probably because marriage is closely related to one’s *happiness* or *unhappiness* and an *unhappy* marriage is more likely to be the topic of the discourse. *Unhappy* has colligations with the article-determiner *an* (760), link verbs, mainly *be* (inclusive of its various forms, henceforth), *feel* (81), *look* (67), *become* (39), prepositions, typically *with* (1892), *about* (765), *at* (424). *Look* seems to indicate that an *unhappy* feeling is often shown outwardly via external manifestations. Its word sketch shows that *unhappy*

is often intensified by amplifiers, like *very* (487), *deeply* (111), *increasingly* (57), *desperately* (54), *extremely* (39) and the emphasiser *really* (51), but less readily combinable with downtoners like *little* (14) or *slightly* (1) because it ‘implies a more ‘intense’ feeling and a ‘stronger’ negative evaluation’ (Wierzbicka, 1999: 63). The nouns used in the ‘pp_with-p’ pattern include *result* (26), *situation* (23), *decision* (23), *life* (20), *outcome* (16), *performance* (14), *plan* (14), *idea* (13) etc., suggesting a semantic component of dissatisfaction. This reveals that the typical cognitive scenario of ‘unhappy’ is that you are not satisfied with what has happened or is happening in a particular situation.

What can be found for *unhappily* in OALECD is shown below:

1. sadly 可悲地; 难过地.
2. unfortunately 不幸地; 不走运地; 遗憾地

From the definitions above, we can see that *unhappily* can translate as *nanguo de, kebei de, buxing de, bu zouyun de, yihan de*. Let us look at *unhappily* (535 hits) in the BOE. Its top collocates show a number of verbs modified by *unhappily*: *marry* (56), *say* (32), *end* (19), *think* (13), *look* (7), *stare* (6) and *sit* (5). If *unhappily* is used to modify a verb of action, it usually means ‘sadly’. If we say someone lives or dies *unhappily*, it probably means his life or death is an unfortunate thing. It is noteworthy that *unhappily*, apart from action verbs, is semantically associated with marriage or the ending of it.

From the 1422 hits of *unhappiness* found in the corpus, an inspection of its top 50 collocates reveals that it has colligations with prepositions, typically *of* (198), *with* (154), *at* (62), *in* (64), *about* (31), *over* (24), possessive adjectives, such as *his* (96), *their* (79), *her*(78), *my* (24), *your*(22), *own* (19), amplifiers, predominantly *deep* (23),

great (17), quality adjectives, like *personal* (16), *public* (11), *human* (7), *marital* (5), and verbs, mainly *express*(11), *cause* (11), *grow* (10), *bring* (7). The close association of the ‘unhappy’ concept with marriage is once more supported by *marital*. *Personal*, *own*, *public* and *express* seem to suggest that *unhappiness* is both extroverted and introverted. *Cause* and *bring* appear to indicate that *unhappiness* has a definite cause.

5.2.2.2 Chinese equivalents of *unhappy*

Next let us move on to the discussion of their Chinese counterparts. *Unhappy** occurs 55 times in the BPC, which yields the following result:

Table 5.2. 1 The most frequent Chinese equivalents of *unhappy** in the BPC

Chinese equivalent	Frequency
<i>buxing</i>	24
<i>bu kuaile</i>	5
<i>bu yukuai</i>	6
<i>bu gaoxing</i>	4
<i>bu xingfu</i>	4
Others	12
Total	55

All of the expressions start with *bu*, which is an adverb preceding adjectives to indicate negation. *Buxing* tops the list because it is the closest equivalent to express *unhappy*’s sense of ‘unfortunate or unlucky; regrettable’. Consider this example:

- 1) 许多在场的人断言,他们在那个不幸 (*buxing*) 的牧师的胸前看到了一个嵌在肉里的红字,与海丝特·白兰所佩戴的十分相似.... out of what had been witnessed on the scaffold. Most of the spectators testified to having seen, on the breast of the *unhappy* minister, a SCARLET LETTER -- the very semblance of that worn by Hester Prynne -- imprinted in the flesh.

In this example, *unhappy* is translated as *buxing*. According to the OED Online, the ‘not happy’ meaning developed from the ‘unlucky’ meaning. In Chinese, there is a separate word ‘*buxing*’, literally meaning ‘not lucky’ or ‘unfortunate’, to convey this meaning.

Unhappy and *unlucky* are two distinct concepts both in Chinese and English, but *unhappy* connects ‘misfortune’ and ‘a feeling of displeasure’ through a conceptual extension, which is missing in Chinese.

In translating, the sense of ‘not happy’ is divided by many Chinese phrases with similar meanings. The four most frequent ones are listed in the table. The last one *bu xingfu* is based on the evaluation of one’s own life. It concerns the satisfaction with different aspects of one’s life, depending more on some cognitive appraisal. It has not only to do with ‘our personal past experience, but also to our social experience. We are always comparing ourselves to others. And we feel good or bad depending on whom we compare ourselves to’ (Myers, 1991: 56). ‘Our social comparisons concern not only money, but also looks, smarts, and various forms of success’ (p.57). Rich people are not necessarily happier than less wealthy people because the former may compare themselves with those above them or have higher expectations toward life. People from different cultures have different criteria to judge life satisfaction. In Chinese, the closest equivalent to this meaning is *xingfu*. Its negative form *bu xingfu* is used to describe life dissatisfaction.

I will discuss the other three expressions meaning ‘not happy’: *bu yukuai*, *bu kuaile* and *bu gaoxing*. *Bu gaoxing* is usually a short-term feeling triggered by a something specific; *bu kuaile* is extended in time and more general; *bu yukuai* is also specific. They can be separated by degree words like *zenme* (to some degree, used in negatives). Consider this example:

- 2) 可凯里·阿什伯恩显得并不怎么高兴 (*gaoxing*) , 因为他明明很不喜欢瑞德,他十分愿意站在米德大夫一边可是又不能说假话。

Carey Ashburn looked *unhappy*, for it was obvious that he, too, disliked Rhett intensely. He gladly would have sided with the doctor but he could not lie.

In this example, *unhappy* is translated as *bu zenme gaoxing*, which means ‘not very happy’. In Chinese, it is very common to insert degree words into negative phrases. To explore the finer distinctions of these Chinese phrases, I will analyse them one by one in detail based on the corpus data.

5.2.2.2.1 *Bu gaoxing*

2561 hits of *bu gaoxing* are found in the CCL corpus and its significant collocates are shown below:

Table 5.2. 2 Significant collocates of *bu gaoxing* in the CCL corpus

Extroversion	Introversion	Modifier	Others
<i>lian</i> 73	<i>xinli</i> 117	<i>hen</i> 407	<i>re</i> 40
<i>manlian</i> 48	<i>xinzhong</i> 28	<i>hao</i> 70	<i>duzi</i> 26
<i>lianshang</i> 29	<i>juede</i> 37	<i>laoda</i> 45	<i>yinwei</i> 66
<i>yangzi</i> 34	<i>gandao</i> 41	<i>feichang</i> 33	<i>weishenme</i> 48
<i>xiande</i> 21		<i>zhen</i> 31	
		<i>ting</i> 21	
		<i>xie/youxie</i> 32/36	
		<i>dian/youdian</i> 41/27	

The first column in the above table lists the collocates which are semantically associated with outward manifestations: *lian/manlian/lianshang* (face/full face/on the face), *yangzi* (appearance) and *xiande* (appear). Sometimes, we could judge whether someone is happy or not based on his or her facial expressions or behaviour. But it is not always so. The collocates in the second column reveal that *bu gaoxing* is also realized as an inner emotional state: *xinli/xinzhong* (in the heart), *xin* (heart) and *juede/gandao* (feel). *Unhappy* also tends to co-occur with *feel* and *look*, as discussed earlier, but not with *in the heart*. Different from English, Chinese tends to add *xinzhong* or *xinli* to emphasise

that what the person experiences is an *unhappy* feeling, but s/he does not necessarily reveal it on the face. Traditionally, Chinese people have followed the philosophy of ‘xi nu bu xing yu se’ (do not reveal your joy and anger on your face). In its culture, it has been taken for granted that people hide their true feelings to others. In such a society, we cannot judge whether a person is happy or unhappy just based on his/her facial expressions. At the same time, in the Chinese discourse, people distinguish clearly an inward feeling from an outward facial expression or bodily manifestation, which might not reflect the true feeling. Consider the following example:

- 3) 人家嘴上不说可心里[不高兴](*bu gaoxing*)
‘He was actually *bu gaoxing*, though he didn’t say so’

Things like this happen quite often. They have to do with the famous Chinese ‘face’ (*mianzi*) issue, which is said to originate from Confucius’ concept of ‘ming’ (名) (status) (Nongren, available at <http://www.my1510.cn/article.php?id=b255454ac109ed9a>). Chinese people take their ‘faces’ very seriously. In speaking or talking, they have to be considerate enough not to hurt other people’s ‘faces’. On the other hand, if someone is praised for his/her achievement, the traditional response would be a denial, though s/he is actually very happy. This is due to the well-known modesty of Chinese people. In the English culture, it is also common to hide the true feelings because it is assumed to be a departure from ‘normal behaviour’ to show feelings ‘over which one has no control’ (Wierzbicka, 1999: 19). Wierzbicka interprets the word ‘emotional’ as a certain unconscious ‘ideology’—the loss of control of one’s feelings would be abnormal. But generally speaking, compared with Chinese, English-speaking people are more direct, frank, and transparent. Another reason to account for such a difference might be the

Chinese specific way to express emotions. Look at the following examples from the

Babel Chinese-English Parallel Corpus¹⁴:

- 4) Xinzhong you youxie nanguo
In the heart but a little bit unhappy
'But he felt *nanguo* about it'
- 5) Xiangzi xinzhong feichang gaoxing
Name in the heart very happy
'Xiangzi was very *gaoxing*'

The above examples are not translated word by word. In English, being happy or unhappy by default is an inner state of emotion, so 'xinzhong' can be left out in the translations.

From the third column in Table 5.2.2, we can see that its modifiers include *hen/laoda/ting/hao/feichang* (very), *zhen* (really) and *xie/youxie/dian/youdian* (a little). In English, adjectives can also be modified by a little (bit). It seems that Both English and Chinese have metaphorised 'small number or amount' as 'to a little degree', which has been referred to as grammaticalisation. The last column lists its other meaningful collocates. Among them, *re* (invite/lead to) is a frequent causative verb to collocate with *bu gaoxing*. It is usually used in the pattern 're mouren *bu gaoxing*' (cause/make someone *unhappy*). *Duzi* (belly) appears in the above table because *yi/man duzi bu gaoxing* (full of *bu gaoxing* in one's belly) is a recurrent pattern, meaning 'very unhappy'. The Chinese culture conceptualises that emotions are located in one's belly. You become *bu gaoxing* because something bad happened and you do not want this, and the event affects your feeling only temporarily. *Wei qu* (grievance/feel wronged) and

¹⁴ This corpus is composed of parallel texts from Chinese to English. It is too small to yield reliable statistical information, so I only use it to retrieve some examples.

shengqi (annoyance) are often associated with *bu gaoxing*, as the following examples illustrate:

- 6) 老师没听解释就批评了我。我感到很委屈 (*weiqu*) , 好几天不高兴 (*bu gaoxing*)。
'The teacher criticised me without listening to my explanation, so I felt very *weiqu*. Because of this I felt *bu gaoxing* for several days.'
- 7) 好像今天不高兴 , 跟谁生气 (*shengqi*) 呢 ?
'You look *bu gaoxing* today. Who are you *shengqi* with?'

If someone is wronged, s/he will feel *bu gaoxing*. If someone is *shengqi*, s/he must be *bu gaoxing*. *Shengqi* is a weak version of anger. XHCBL glosses it as 'yin buhe xinyi er *bu yukuai*', which literally means 'not happy due to dissatisfaction'. Bad events---a rejection, failing in an exam, losing a game---may make you *bu gaoxing*. But our moods will return to normal soon, because 'we humans have an enormous capacity to adapt to changed circumstance' (Myers, 1991: 48). In some cases, the feeling of *bu gaoxing* may linger longer, but it usually lasts for a short time because one's current mood is more affected by what is happening. Thus we can accept easily the past bad events and adjust our moods to welcome something new.

In the 2nd sense of OALECD, *unhappy* is glossed as 'feeling worried or annoyed because you do not like what is happening in a particular situation'. While *unhappy* and *bu gaoxing* are similar in being annoyed, *bu gaoxing* has nothing to do with worry. Like *unhappy*, *bu gaoxing* is a feeling when something I do not like happened. It is quite specific. In other words, *bu gaoxing* has a definite cause, as indicated in its collocates *yinwei* (66/because) and *weishenme* (48/why). Further, a simple glance of its concordance reveals that *bu gaoxing* is overwhelmingly used predicatively.

5.2.2.2.2 *Bu yukuai*

Bu yukuai occurs 1408 times in the corpus. Below are its significant collocates:

Table 5.2. 3 Significant collocates of *bu yukuai* in the CCL corpus

Introversion	Modifier	Verb	Noun	Others
<i>xinli</i> 40	<i>hen</i> 163	<i>fasheng</i> 89	<i>shiqing</i> 133	<i>meiyou</i> 37
<i>xinzhong</i> 16	<i>ji</i> 38	<i>chansheng</i> 25	<i>shi</i> 107	<i>zhijian</i> 22
<i>xinqing</i> 46	<i>feichang</i> 27	<i>nao</i> 26	<i>shijian</i> 24	
<i>qingxu</i> 36	<i>yidian</i> 18	<i>yinqi</i> 25	<i>jingli</i> 21	
<i>gandao</i> 80		<i>yudao</i> 23	<i>guoqu</i> 22	
<i>ganjue</i> 33		<i>shi</i> 103	<i>shenghuo</i> 31	
<i>juede</i> 37		<i>ling</i> 99	<i>gongzuo</i> 15	
		<i>rang</i> 40		

The table shows that, semantically, *bu yukuai* has a preference for words related to internal states, including sensational verbs, predominantly *gandao/ganjue/juede* (feel), prepositions indicating the assumed loci of emotions, typically *xinli /xinzhong* (in the heart), and its superordinates, such as *xinqing* (frame of mind), *qingxu* (mood). Look at the following instance, which clearly shows that *bu yukuai* is a frame of mind or mood:

- 8) 他跳槽到了广东某地的一家外资企业，工资比康惠高了近一倍，
但心情并[不愉快] (*bu yukuai*)
‘He changed to another foreign enterprise, with his pay doubled, but
actually he was *bu yukuai*’

Hence, it may be argued that *bu yukuai* is more realised as an internal emotional state, focusing mostly on what is mentally experienced. It can also be seen, from the table, that *bu yukuai* colligates frequently with intensifiers, typically *hen* (very), *ji* (extremely), *feichang* (very much) and *yidian* (slightly). It has a semantic preference for things which might make people *bu yukuai*, such as *shiqing/shi/shijian* (thing/event), *jingli* (experience), *guoqu* (past), and aspects in terms of which *bu yukuai* might happen, typically *shenghuo* (life), *gongzuo* (work), as exemplified in the following example:

- 9) 日常生活中处处都可能遇到 [不愉快] (*bu yukuai*) 的事
‘In everyday life, it is likely to encounter *bu yukuai* things’.

It seems that *bu yukuai* is often used as an attributive, which is different from *bu gaoxing*, which is more likely to be used as a predicate. Another difference is that ‘*meiyou bu yukuai*’ means ‘there is nothing unhappy’, where *meiyou* is the same as ‘no’ or ‘not have’, while ‘*meiyou bu gaoxing*’ means ‘not unhappy’, where *meiyou* is used to indicate negation. Further, *bu yukuai* often occurs between two parties, so *zhijian* (between) is salient in its collocates. Examples are shown below:

- 10) 我不想因为这件事情而使我们之间(*zhijian*)发生(*fasheng*)[不愉快] (*bu yukuai*) ‘I don’t want any *bu yukuai* to happen between us because of this’
- 11) 所以我们之间(*zhijian*)并没有(*meiyou*)什么[不愉快] (*bu yukuai*),大家还是好离好散 ‘So there was no *bu yukuai* between us. We broke up peacefully’

The phrase *bu yukuai* could be used as a noun, meaning something which makes people *bu yukuai*, just as shown above. Except *fasheng* (occur) and *meiyou* (no), it has some other typical verbal collocates:

- 12) 人们必须保证不要出现(*chuxian*)令人震惊的[不愉快] (*bu yukuai*) ‘People must make sure that nothing shockingly *unhappy* occurs’
- 13) 洛宾一点也没有察觉三毛因为拍电视而引起(*yinqi*)的[不愉快] (*bu yukuai*) ‘Robin was not at all aware of Sanmao’s sadness caused by TV show making’

It should be noted that, in Chinese, *fasheng/chuxian/chansheng* (occur/emerge), like *yinqi* (cause), are all transitive verbs. *Nao* is a causative verb which often collocates with *bu yukuai*, meaning ‘leading to something unpleasant’:

- 14) 闹 得 很 [不愉快] (*bu yukuai*) ‘led to something very *bu yukuai*’

As the table shows, *bu yukuai* frequently co-occurs with *shi/ling/rang* (make/cause), suggesting a semantic preference for causative words. It is semantically associated with disappointment, as exemplified below:

- 15) 正是事先预料的那样, 是令人失望 (*ling ren shiwang*) 和极[不愉快] (*bu yukuai*) 的 ‘As expected, it was disappointing and extremely *bu yukuai*’

It seems odd to say ‘someone is disappointed and very *bu gaoxing*’, because ‘disappointed’ implies ‘*bu gaoxing*’, so it’s completely redundant to add it. But disappointment can be juxtaposed with *bu yukuai* since they each emphasises a different aspect. *Bu yukuai* is closer to ‘unpleasant’ in feeling as both give no pleasure to the mind. *Bu yukuai* is more intense than *bu gaoxing*. I will further explore the subtle differences between *bu gaoxing* and *bu yukuai* by looking at their causes. Look at these examples:

- 16) 然而生活中存在的个别不文明行为却使多数同学感到[不愉快] (*bu yukuai*) ‘But any uncivilised behaviour in everyday life makes many students feel *bu yukuai*’
- 17) 大部分人闻到这味道就感到[不愉快] (*bu yukuai*) ‘most people feel *bu yukuai* when smelling this’

It can be seen from these examples that a bad smell or uncivilised behaviour can make people *bu yukuai*. *Bu yukuai* might mean people feel upset or uncomfortable. One may feel *bu yukuai* when s/he has worries or burdens. The feeling of *bu yukuai* can last long or be triggered again even long after the unpleasant thing. However, one feels *bu gaoxing* because what happened is not what s/he wants. *Bu gaoxing* is quite specific and short. It may disappear soon because the situation changes or something happy happens afterwards.

5.2.2.2.3 *Bu kuaile*

The CCL corpus only generates 255 instances of *bu kuaile*, whose significant collocates are shown below:

Table 5.2. 4 Significant collocates of *bu kuaile* in the CCL corpus

Modifier	Introversion	Others
<i>hen</i> 19	<i>xinli</i> 6	<i>yinwei</i> 11
<i>feichang</i> 17	<i>gandao</i> 8	<i>weishenme</i> 8
<i>yidian</i> 7	<i>juede</i> 6	<i>yuanyin</i> 7

The table shows that it has a colligation with intensifiers: *hen* (very), *feichang* (very much) and *yidian* (a little). Its semantic preference for words of introversion, characteristically *gandao/juede* (feel), *xinli* (in the heart), shows that *bu kuaile* is more experienced as an inner state. The last column reveals that the reasons why people are *bu kuaile* are often discussed in the discourse, suggesting that *bu kuaile* is more specific because it has a definite reason, as indicated in *yinwei* (because), *weishenme* (why) and *yuanyin* (reason). To determine what *bu kuaile* exactly means, we have to look at some examples:

- 18) 结婚是人生最美满快乐的事，我和我内人都是个中人，假使结婚不快乐(*bu kuaile*)，我们应该苦劝两位别结婚 ... Marriage is the happiest event of one's life. My wife and I both know what it is. If marriage *weren't happy*, we should be exhorting you not to marry.
- 19) 就因为张蕊玲的阴谋，我们就要永远的[不快乐] (*bu kuaile*) 吗？
‘Are we going to be *bu kuaile* for good, just because of Zhang Ruiling’s conspiracy?’

The above examples seem to suggest that *bu kuaile* lasts long since it is associated with marriage and permanency. More examples:

- 20) 有钱的人，通常并不[不快乐] (*bu kuaile*)
‘Usually, rich people are actually *bu kuaile*’
- 21) 凡是童年[不快乐] (*bu kuaile*) 的人都特别脆弱
‘People with a *bu kuaile* childhood are especially vulnerable’
- 22) ‘上班族’[不快乐] (*bu kuaile*) 的原因是压力太大,压力主要来自工作、经济因素和家庭 ‘The reason why ‘shangbanzu’ are *bu kuaile* is because they have much pressure from work, financial matters and family’
- 23) 但我的心中充满隐忧，我其实是极度地[不快乐] (*bu kuaile*) 的
‘But I have a lot of worries. I’m actually extremely *bu kuaile*’
- 24) 有时候这种反省会把我弄得非常[不快乐] (*bu kuaile*)，惊讶的发现自己的失误竟是那么多 ‘Sometimes the reflection made me very *bu*

kuaille, because I found that I had made a surprisingly large number of mistakes’

Rich people are *bu kuaille* probably because they want more than money can provide. A *bu kuaille* childhood is often associated with an abnormal life. Usually a child enjoys a carefree life under the care of his/her parents. A *bu kuaille* childhood often means the child has to live a life with worries. When we say ‘a *bu kuaille* de childhood’, we mean that we divide the life into several periods, one of which—childhood—is *bu kuaille*. It seems that we can limit it to a certain period of time and then judge whether someone is *kuaille* or *bu kuaille* in this period. ‘Shangbanzu’ in Example 22 are *bu kuaille* because of their pressure, suggesting that worries lead to *bu kuaille*. The reflection in Example 23 made me *bu kuaille* because I found I had made much more mistakes than expected. This might make me think more about what caused these mistakes. *Bu kuaille*, unlike *bu gaoxing*, often makes people think more about what happened or what it may further lead to. In other words, one feels *bu kuaille* because s/he has something to worry about in life. A carefree person would be *kuaille* even s/he is not rich. In this sense, it is related to life dissatisfaction. We say ‘shengri (birthday) *kuaille*’ to others because we wish the addressee, at least, a carefree and joyful day. *Bu gaoxing* is an emotional, and more often a physical reaction as well, to a specific thing based on the judgement that what happened is not what s/he wants. People who have worries or cannot forget their worries are unlikely to enjoy life, so they will be *bu kuaille*.

5.2.2.3 Comparison

Now let us come to the comparison between *unhappy* and its Chinese translation equivalents. One difference not relevant to meaning analysis is that *bu yukuai* can also

be used as a noun. Another notable difference among *bu gaoxing*, *bu yukuai* and *bu kuaile* would be that *bu yukuai* can be used to modify a thing, place or situation that makes people unhappy, such as *shi* (event), *changhe* (occasion/situation), *qifen* (atmosphere), *jingli* (experience), *jiyi* (memory) and *hunyin* (marriage), as shown in the table below:

Table 5.2. 5 Comparison of some collocates of *bu gaoxing*, *bu yukuai* & *bu kuaile*

	<i>shi</i>	<i>changhe</i>	<i>qifen</i>	<i>jingli</i>	<i>jiyi</i>	<i>hunyin</i>
<i>bu gaoxing</i>	18	0	0	0	0	0
<i>bu yukuai</i>	219	5	5	12	7	7
<i>bu kuaile</i>	2	0	0	0	0	0

In this sense, it is more similar to *unhappy*. *Bu gaoxing* and *bu kuaile* can only co-occur with thing (event), but not with the others. *Lingren/rangren* (make/cause) should be added to them to express the making-unhappy meaning. *Shi* (event) seems to be very special in Chinese in that it does not follow the common rules and can collocate with many words. So I will exclude it from the analysis. These expressions also differ in many other aspects. Table 5.2.6 summarises their respective features:

Table 5.2. 6 Features of unhappy expressions

	<i>unhappy</i>	<i>bu gaoxing</i>	<i>bu yukuai</i>	<i>bu kuaile</i>	<i>bu xingfu</i>
Extroverted	√	√			
General	√		√	√	√
Specific	√	√	√		
Length	short or long	usually short	short or long	usually long	long
Making unhappy	√		√		
Annoyed	√	√	√		
Worried	√			√	
Unfortunate	√				
Not suitable	√				

It can be seen from the table that *unhappy* seems to be a good-for-anything word which covers the meanings of all four Chinese equivalents. I will focus on the semantic differences caused by cultural specificity. Chinese distinguishes these expressions in terms of the dimensions of extroversion, specificity, generality and length. No Chinese word has all of these features. *Unhappy* and its translational equivalents are seemingly semantically similar, but if you examine them individually, they will exhibit distinct features. In other words, they have different functional loads. *Unhappy* covers a wider semantic range, also meaning ‘making unhappy’, ‘annoyed’, ‘worried’, ‘unfortunate’ and ‘unsuitable’ etc. Interestingly, some Chinese words also have some of these senses, like *bu yukuai* meaning ‘making unhappy’ or ‘annoyed’, *bu gaoxing* meaning *shengqi*, and *bu kuaile* implying ‘worried’. However, the meanings of ‘unlucky or unfortunate’ and ‘not suitable’ have nothing to do with the Chinese expressions about unhappy feelings.

5.2.3 *Sad** and their Chinese equivalents

5.2.3.1 *Sad**

Sad is a generic term for unhappy feelings. First, let us look at the senses from the three chosen English dictionaries:

LDCE:

▣ **【FEELING UNHAPPY】** unhappy, especially because something unpleasant has happened ▣ **【MAKING YOU UNHAPPY】** a sad event, situation etc makes you feel unhappy

▣ **【NOT SATISFACTORY】** very bad or unacceptable ▣

【LONELY】 a *sad* person has a dull, unhappy, or lonely life

▣ **【BORING】** *informal* boring or not deserving any respect

OALECD:

① showing or causing sorrow; unhappy 悲哀的; 忧愁的; 难过的 ②
 [attrib 作定语] worthy of blame or criticism; bad 该受责备或批评的;
 坏的 ③ making one feel pity or regret 令人遗憾或惋惜的 *sadness* *n* 1
 [U] being sad 悲哀; 忧伤. 2 [C usu *pl* 通常作复数] thing that makes
 one sad 令人悲哀或忧伤的事物

COBUILD:

1 [ADJ] oft ADJ that/to-inf, ADJ about *n* If you are **sad**, you feel unhappy, usually because something has happened that you do not like.
 ≠ happy

2 [ADJ] usu ADJ *n* **Sad** stories and **sad** news make you feel sad.

3 [ADJ] A **sad** event or situation is unfortunate or undesirable.

4 [ADJ] usu ADJ *n* If you describe someone as **sad**, you do not have any respect for them and think their behaviour or ideas are ridiculous.

(INFORMAL) = pathetic

A simple glance of the definitions above will show that these three dictionaries cannot agree about the numbers and meanings of the senses *sad* has. A careful analysis of their senses reveals that, basically, they agree about two senses: a) feeling unhappy and b) making someone unhappy. OALECD and LDCE share the sense of ‘bad’. LDCE and COBUILD agree about the meaning of ‘undesirable’. Both OALD and COBUILD say *sad* is associated with ‘unfortunate’ in some contexts. All of these meanings can be seen as extensions of its original meaning. The meaning of ‘making unhappy’ may be associated with numerous things; in other words, *sad* can be used to talk about a large number of things that make someone unhappy. However, in discourse, it has its preferences to describe. The most frequent ones have attained the attention of lexicographers. However, they do not agree about classifying criteria, so they produce different number of senses and define them in different ways.

Now let us look at the data for *sad* (18212 hits) and *sadness* (4161 hits) from the BOE. From their top 50 collocates, we can roughly set out the components of *sad* and *sadness* as follows:

Core: *sad*

Colligations:

- (a) article-determiners: *a* (2689), *the* (2073)
- (b) link verbs: *be* (3009), *feel* (428), *look* (194)¹⁵
- (c) prepositions: *about* (270), *for* (426), *to* (1158)
- (d) modifiers: *very* (1612), *really* (318), *little* (230), *rather* (128), *terribly* (90), *extremely* (69)

Semantic preferences:

- (a) nouns modified by *sad* (making unhappy): *day* (592), *thing* (418), *story* (243), *fact* (225), *news* (175), *truth* (155), *part* (128), *state* (123), *time* (116), *case* (105), *song* (89), *tale* (85), *reality* (76), *loss* (75), *end* (87), *commentary* (65)
- (b) facial parts or expressions (showing sadness): *eye* (89), *face* (80), *smile* (68)

Core: *sadness*

Colligations:

- (a) article-determiner: *the* (490)
- (b) possessive adjectives: *her* (79), *his* (68), *their* (42), *my* (35), *your* (20), *our* (18)
- (c) prepositions: *of* (795), *with* (245), *in* (237), *at* (142), *for* (133), *about* (61), *over* (37)
- (d) modifiers: *great* (213), *deep* (80), *profound* (39), *real* (31), *terrible* (28), *overwhelming* (20), *huge* (12), *intense* (11), *immense* (10), *extreme* (10), *some* (50), *much* (34), *little* (20)

Semantic preference:

verbs of feeling or expressing: *express* (46), *feel* (25)

The frequent co-occurrence of *sad* with both *feel* and *look*, as well as the semantic preference of sadness for *express* and *feel*, seems to suggest that *sad* is both extroverted and introverted. The corpus data presented above show that, different from *unhappy*, which is often intensified by amplifiers but not downtoners, *sad* frequently collocates with *(a) little* (230). This is consistent with Wierzbicka's claim that *unhappy* is conceived as a more 'intense' feeling and a 'stronger' negative evaluation than *sad*

¹⁵ The number also includes its frequency as a noun, though very small; henceforth.

(Wierzbicka, 1999: 63). Speaking of their difference, Wierzbicka makes the following comments (1999: 63):

... *unhappy* –in contrast to *sad*--does not suggest a resigned state of mind. If in the case of *sadness* the experiencer focuses on the thought ‘I can’t do anything about it’, in the case of *unhappiness* he/she focuses on some thwarted desires (‘I wanted things like this not to happen to me’), and hence it is more closely associated semantically with *unhappy*. The attitude is not exactly ‘active’ because one doesn’t necessarily want anything to happen, but it is not ‘passive’ either, for one doesn’t take the perspective ‘I can’t do anything about it’.

According to Wierzbicka, the experiencer of *sadness* is quite passive due to helplessness, while the experiencer of *unhappiness* is much less so because its focus is on ‘some thwarted desires’. Let us look at their significant collocates in the ‘and/or’ pattern to see whether *unhappy* and *sad* differ in resignation:

Sad: *lonely* (109), *disappointed* (25), *regrettable* (17), *pathetic* (26)

Sadness: *regret* (32), *disappointment* (24), *loneliness* (15), *hopelessness* (11), *pity* (11), *vulnerability* (8), *self-pity* (5)

Unhappy: *lonely* (32), *confused* (23), *depressed* (18), *insecure* (17), *anxious* (16), *angry* (16), *tense* (14), *uncomfortable* (13)

Unhappiness: *frustration* (13), *stress* (10), *depression* (8), *pain* (8), *misery* (6), *anxiety* (6), *loneliness* (6)

The data presented above indicate that *sadness* is closely associated with *regret*, *disappointment*, *loneliness*, *hopelessness*, *pity*, *self-pity*, and *vulnerability*. It can be inferred that the core of *sadness* is that ‘I am disappointed and hopeless about what happened and I wish things could happen differently’. This is in line with Wierzbicka’s claim that *sadness* implies being more passive. It is also true that the collocates for *unhappy* and *unhappiness* above do not show any apparent resignation. Nonetheless, as mentioned in the previous section, *unhappy* has a marked tendency to occur with *result*, *situation*, *decision*, *life*, *outcome*, *performance*, *plan*, *idea* etc., implying a semantic core

of dissatisfaction. This shows that the typical cognitive scenario of ‘unhappy’ would be that ‘this is not what I want and I am not satisfied with this’. In other words, the main differences between these two concepts lie in the contrast between helplessness and dissatisfaction. Further, as presented earlier, the nouns modified by *sad* are usually short or specific, such as *day, thing, story, fact, news, truth, case, song, tale*, or facial parts or expressions, like *eye, face, smile*. However, *unhappy* tends to modify *marriage, childhood, family, life, relationship*. We may speak of *unhappiness* of a moment or of something rather specific, but we often reserve the term for longer periods or something more general, which often needs to be measured or judged from a certain perspective.

Now let us move on to examine *sadly*. Below are its dictionary entries:

OALECD:

1 in a *sad* manner 悲哀地; 忧愁地 **2** regrettably 令人遗憾地; 惋惜地: *a sadly neglected garden* 荒废得很可惜的花园.

3 unfortunately 不幸地

LDCE:

▣ in a way that shows that you are *sad* = unhappily

▣ [sentence adverb] unfortunately

▣ very much - used when talking about bad situations or states

The garden's been sadly neglected.

It can be seen that LDCE lacks the sense of ‘regrettably’, but it includes the sense of ‘very much’. Their examples show that the two dictionaries interpret ‘a sadly neglected garden’ in different ways. This is quite common in lexicography. They are only interpretations of lexicographers, rather than the absolute truth. Now we come to its corpus data. 8178 hits of *sadly* occur in the corpus. An examination of its top 50 collocates shows that it has a strong colligation with verbs. On the one hand, *sadly* has a semantic preference for verbs to do with physical manifestations, such as (*shake one's*)

head (197), *smile* (140), *nod* (26), *stare* (22), *seem* (19), suggesting that *sad* is very much outwardly shown via facial expressions or bodily events. On the other hand, it is semantically associated with a group of words which are related to bad situations or states, like *miss*/235 (sadly missed), *lack*/97 (sadly lacking), *mistake*/48 (sadly mistaken), *overlook*/29 (sadly overlooked), *neglect*/28 (sadly neglected), *disappoint*/19 (sadly disappointed) etc., where *sadly* can be interpreted either as ‘regrettably’ or ‘very much’. For its sense of ‘unfortunately’, *die sadly* is a frequent collocation. In addition, its modifiers include *rather* (31), *very* (37), *(a) little* (27).

Sadden is glossed as ‘(cause sb to) become sad (使某人) 悲哀, 忧愁 [to make someone *beiai* or *youchou*]¹⁶ in OALECD. Actually, I do not think *youchou* is the main component of *sadden*, but this is not the focus of my research. Its top collocates, based on 1785 hits of *sadden*, show that it colligates strongly with *be* (550), the preposition *by* (521), and modifiers, typically *deeply* (189), *very* (54), *greatly* (21), *really* (20), *profoundly* (11), *extremely* (9), *particularly* (7), *genuinely* (5), *terribly* (5), *little* (5). Its concordance reveals that the causes for being *saddened* include *passing*, *death*, *loss*, *departure*, *news*, *verdict*, *accident*, *decision*, *tragic events/tragedy* etc. Its ‘and/or’ word sketch tells us that being *saddened* tends to occur after being *shocked* (112) and/or *disappointed* (24) and/or being *sickened* (11). Most of us could manage to regain our equilibrium soon and move on.

5.2.3.2 Chinese equivalents of *sad* *

¹⁶ Throughout the thesis, the words in square brackets are added by the author to explain the Chinese definitions in the dictionaries.

133 instances of *sad** occur in the BPC and their most frequent Chinese equivalents are shown as follows:

Table 5.3. 1 The most frequent Chinese equivalents of *sad** in the BPC

Chinese translational equivalent	Frequency
(<i>ling ren /rang ren/ jiao ren</i>) <i>shangxin</i>	39
(<i>ling ren</i>) <i>beishang</i>	37
<i>beican</i>	11
<i>youshang</i>	10
<i>beiai</i>	10
(<i>rang ren</i>) <i>nanguo</i>	8
Others	18
Total:	133

In English, *sad* can be used to describe not only people, expressions, feelings, but also things and situations causing sadness. However, as mentioned earlier, Chinese emotion words have to be preceded by ‘lingren/jiaoren/rangren/shiren’(to make/cause) to express causative meanings. There are specific Chinese words which are used to describe different things and situations. For example, *beican* is used to talk about a miserable situation or bitter experience, as illustrated in the following example:

- 1) You think that her father, even in this *sad* state, will submit himself to her; do you not? 你相信她的父亲即使在目前这种 **悲惨**(*beican*)的状况下也会服从她么？

Consider more examples from the parallel corpus:

- 2) Then she had been pale and *sad* but there had been buoyancy about her. 那时她尽管**面黄瘦**(*mian huang shou*) , 但还显得比较轻松活泼。
- 3) ‘Oh, ‘it’s only--about my own self,’ she said, with a frail **laugh of sadness**, fitfully beginning to peel ‘a lady’ meanwhile. ‘哦 , 这只是——关于我自己的事,’ 她说完 , **苦笑** (*kuxiao*)了一下, 同时又断断续续地动手把 ‘夫人’ 的花蕾剥开。
- 4) It was a *sad* sight, but it showed us that the anchorage was calm. 这是幅 **凄凉** (*qiliang*)的景象 , 但这也告诉我们这锚地非常平静。
- 5) In the *sad* moonlight, she clasped him by the neck, and lad her face upon his breast. 她在**凄清**(*qiqing*)的月光下搂住了爸爸的脖子,把脸靠在他的胸脯上。

- 6) But she can be so cruel and it comes so suddenly and such birds that fly, dipping and hunting, with their small *sad* voices are made too delicately for the sea. 然而她能变得这样残暴，又是来得这样突然，而这些飞翔的鸟儿，从空中落下觅食，发出细微的哀鸣(*aiming*)，却生来就柔弱得不适宜在海上生活。
- 7) Her face was puckered in the *sad* bewilderment of an old ape but there was determination in her jaw. 她那张皱脸孔，像只惶惑不安的老猴似的，不过那下颚却说明她心中早已打定了主意。
- 8) But he made a *sad* business of it with his unsteady hand, and a smothered titter rippled over the house. 可他的手不听使唤，结果把图画得不象样，引得大家暗地里忍俊不禁。
- 9) The piano was *sadly* out of tune but some of the chords were musical... 钢琴是严重(*yanzhong*)走调了，但有的和弦听起来仍然很美...
- 10) In their *sad* doubts as to whether their son had himself any right whatever to the title he claimed for the unknown young woman, Mr and Mrs Clare began to feel it as an advantage not to be overlooked that she at least was sound in her views; 克莱尔先生和克莱尔太太很有些(*hen youxie*)怀疑他们的儿子声明那个他们不认识的年轻姑娘拥有的资格，他们的儿子是不是就有权利得到他说的那种资格，他们开始觉得有一个不能忽视的优点，那就是他的见解至少是正确的。
- 11) Behind those doors lay the beauty of the old days, and a *sad* hunger for them welled up within her. 那些门背后藏着往日的美好，而现在她心里正苦苦(*kuku*)渴望着重新见到它。

If *sad* is associated with a countenance, the Chinese translation would often be *huangshou* (yellow and thin), as shown in (2). If it is used to modify a smile or laugh, the Chinese translation will be *kuxiao* (*lit.* bitter laugh/smile), as exemplified in (3). In (4), (5) and (6), *sad* was translated as *qiliang* in *a sad sight*, *qiqing* in *sad moonlight*, and *aiming* in *sad voices*. Chinese has specific words to describe these things. *Sad* can collocate with numerous words from a wide range of domains, concrete or abstract, such as *bewilderment* and *business* shown in (7) and (8). It seems that *sadly* can also be used to indicate a degree, as exemplified in (9). Example (10) and (11) show that *sad* also has such a function, which is missing in dictionary descriptions.

I have discussed *beishang* and *beiai* in Section 5.2.1, so here I will only analyse the other three Chinese equivalents of *sad*: *shangxin*, *nanguo* and *youshang*. *Shangxin* occurs 39 times in the parallel corpus and it is significantly higher than the frequency of any other other equivalent shown in the table. It seems that *shangxin* is the commonest Chinese word to describe a *sad* feeling. Literally, it means the heart is hurt, which is reminiscent of the English *heart-broken*. But they actually have different meanings, to which I will return in Section 5.2.4. In Chinese, many descriptions of physical pains can be used to depict inner emotional states. Packard (2000) referred to this as component metaphorical lexicalisation in which one or both of the individual word components take on a metaphorical meaning, while the overall meaning of the word continues to be a compositional sum of the meanings of its metaphorical parts. *Shang* and *xin* here have taken on metaphorical meanings, though their relationship remains the same, resulting in *shangxin* acquiring the meaning of a *sad* feeling. But what does *shangxin* exactly mean? XHCBL provides the following:

由于 遭受 不 如意 的 事情 而 心里 痛苦
youyu zaoshou bu ruyi de shiqing er xinli tongku:
 because undergo not as wished ART thing CONJ in the heart pain
 ‘suffering mentally due to an unpleasant experience’

Tongku is found in the definition for *shangxin*. Then what does *tongku* mean? Its definition from XHCBL is shown as follows:

身体 或 精神 感到 非常 难受
shenti huo jingshen gandao feichang nanshou
 body or mind feel very uncomfortable
 ‘feeling very uncomfortable physically or mentally’

Another word *nanshou* is introduced. According to XHCBL, *nanshou* is:

Shangxin; bu tongkuai 不痛快
shangxin not happily relieved of a burden/overjoyed

It seems that we will never get out of the circle of cross-explanation if we restrict ourselves to dictionaries. So a corpus search is necessary.

5.2.3.2.1 *Shangxin*

Below are the significant collocates of *shangxin* (4219 hits) from the CCL corpus:

Table 5.3. 2 Significant collocates of *shangxin* in the CCL corpus

Introversion/ extroversion	Cause causative verb	+ Modifier	Feeling	Others
<i>gandao</i> 143	<i>yinwei</i> 90	<i>hen</i> 338	<i>tongku</i> 55	<i>xiangqi</i> 24
<i>ganjue</i> 13	<i>yin</i> 29	<i>feichang</i> 78	<i>haipa</i> 18	<i>shi</i> 232
<i>juede</i> 60	<i>si/siqu</i> 95/15	<i>shifen</i> 65	<i>shiwang</i> 39	<i>guoqu</i> 27
<i>xinli/xinzhong</i> 17/12	<i>likai</i> 28	<i>tebie</i> 20	<i>fennu</i> 29	<i>wangshi</i> 24
<i>yangzi</i> 30	<i>shiqu</i> 18	<i>shizai</i> 26	<i>shengqi</i> 30	<i>gushi</i> 18
<i>kan/kandao</i> 87/27	<i>buxing</i> 15	<i>jile</i> 45	<i>qi</i> 37	<i>anwei</i> 20
<i>lianshang/lian</i> 16/13	<i>shi</i> 223	<i>ji</i> 33	<i>nanguo</i> 23	<i>quan</i> 18
<i>ku/kuqi</i> 436/33	<i>jiao</i> 127	<i>guodu</i> 17	<i>juewang</i> 22	<i>chu</i> 53
<i>lei/yanlei</i> 75/72	<i>ling</i> 113	<i>buyi</i> 18	<i>weiqu</i> 21	<i>ci</i> 42
<i>leishui</i> 18	<i>lingren</i> 33	<i>yidian</i> 19		<i>yizhen</i> 19
<i>luolei/liulei</i> 74/27	<i>re</i> 29	<i>yixie</i> 15		<i>ziji</i> 139
<i>tongku</i> 35				

From the first column in the table, it can be clearly seen that *shangxin* is something one feels (*gandao/ganjue/juede*) in the heart (*xinli/xinzhong*). *Shangxin* is also an emotion involving external manifestations, as implied in its collocates: *lianshang/lian* (on the face/face), *yangzi* (appearance), *kan/kandao* (see), *ku* (cry), *kuqi* (weep), *lei/yanlei/leishui* (tears), *luolei/liulei* (shed tears) and *tongku* (cry bitterly). The second column shows that *shangxin* co-occurs frequently with *yinwei/yin* (because), suggesting that it is an emotion with a clear reason, such as death (*si/siqu*), departure (*likai*), loss (*shiqu*) and misfortune (*buxing*). It also often collocates with causative verbs, like *shi/jiao/ling* (*ren*) (cause/make), *re* (invite/lead to). The third column gives its modifiers, typically *hen/feichang/shifen* (very), *tebie* (particularly), *shizai* (really), *guodu* (exceedingly), *ji/jile* (extremely), *buyi* (endlessly), *yixie/yidian* (a little).

From the fourth column, it can be seen that, like *sadness*, *shangxin* has various negative feelings in its company, such as *tongku* (anguish), *haipa* (fear), *shiwang* (disappointment), *fennu* (indignation), *qi/shengqi* (annoyance), *nanguo* (sadness), *juewang* (despair) and grievance (*weiqiu*). The last column lists its other collocates which can contribute to disclosing its meaning. *Xiangqi* (think of) might suggest that *shangxin* is a feeling stirred up by thinking of something *sad*. Usually *shangxin* is not used to describe things or situations which make people *shangxin*. It can be preceded by causative verbs to fulfil such a function. However, it has some exceptions. Some specific things can be modified by *shangxin*, such as *shi* (thing), *wangshi* (past events), *guoqu* (past) and *gushi* (story). One's *shangxin* could be lessened or overcome after being consoled (*anwei*) or persuaded (*quan*). *Shangxin chu* (place) means what exactly makes someone *shangxin*. This is Chinese specific. Chinese tends to use *chu* or *dian* (point) to highlight what really matters. Another Chinese specific feature is that it tends to use *ci* (number of times) to indicate the frequency of *shangxin*. In other words, *shangxin* in Chinese is conceptualised as repeated happenings. *Shangxin* could last for a short period (*yizhen*). In addition, one often feels *shangxin* on his/her own (*ziji*), not wanting others to know his/her true feelings. From *shangxin*'s external manifestations, such as *ku*, we may infer that *shangxin* is prototypically a transient emotion since you cannot continue crying for a long time, but a common Chinese phrase shows that *shangxin* can last long:

- 12) *zhengri yi lei xi mian*
 all day long with tears wash face
 'so sad that the face is always covered with tears (possibly for even several months)'

You may feel *shangxin* when you are reading a *sad* story, when you think of a *sad* past event, or because your dog has died, because you have been cheated by one of the persons you trusted most, because the person you have been loving dearly fell in love with one of your best friends, because you are ignored by your teachers at school, because, if you are young, other pupils have been picked up by their parents and you know no one will come to collect you, because your parents divorced without considering what you might feel, or because you have heard one of your close relatives or friends has died recently, etc. It may be argued that *shangxin* can be elicited by anything which will make you produce tears or over which you want to since there is an old Chinese saying:

- 13) *naner you lei bu qing tan*
 fine man has tears not lightly shed
 ‘A real man won’t shed tears lightly’.

5.2.3.2.1 *Nanguo*

Another similar term to describe sadness is *nanguo*. The internal structure of this disyllabic lexeme is ‘the adjective prefix *nan+guo*’ difficult + live. Similar examples of this structure are:

- nande* difficult+get ‘hard to come by’
nanchi difficult+eat ‘taste bad’
nanshuo difficult+say ‘difficult to say’

A common feature of this group of words is simple and less formal. Below is the definition for *nanguo* found in XHCBL:

1. *bu rongyi guohuo* (be hard off; have a hard time)
2. *the same as nanshou* (*shangxin; bu tongkuai*)

Nanshou is a weaker and even less formal version of *nanguo*. Let us look at the corpus data. The significant collocates of *nanguo* (4152 hits) are shown below:

Table 5.3. 3 Significant collocates of *nanguo* in the CCL corpus

Introversion/ extroversion	Modifier	Feeling	Others
<i>xinli</i> 622	<i>hen</i> 642	<i>shangxin</i> 45	<i>xiangdao</i> 19
<i>xinzhong</i> 152	<i>feichang</i> 208	<i>tongku</i> 22	<i>xiangqi</i> 17
<i>xinqing</i> 36	<i>wanfen</i> 12	<i>shengqi</i> 16	<i>shi</i> 185
<i>gandao</i> 355	<i>tebie</i> 34	<i>beishang</i> 14	<i>rang</i> 120
<i>juede</i> 146	<i>ji/jile</i> 22/82	<i>weiqu</i> 14	<i>jiao</i> 87
<i>ganjue</i> 17	<i>youxie</i> 21	<i>zhenjing</i> 13	<i>ling</i> 55
<i>yangzi</i> 55	<i>yixie</i> 18	<i>shiwang</i> 13	<i>ci</i> 6+6
<i>lian</i> 15	<i>yidian</i> 27	<i>haipa</i> 13	<i>yizhen</i> 53
<i>kandao</i> 108	<i>youdian</i> 21	<i>houhui</i> 12	<i>anwei</i> 25
<i>kanjian</i> 22		<i>xiukui</i> 12	<i>quan</i> 21
<i>ku</i> 69			<i>rizi</i> 157
<i>lei</i> 18			<i>shiqu</i> 15
<i>lúlei</i> 12			<i>bing</i> 15
			<i>qushi</i> 13
			<i>ziji</i> 176

Our discussion will start with the second column showing its modifiers: *hen/feichang* (very), *tebie* (particularly), *wanfen/ji/jile* (extremely), *youxie/yixie/ yidian/youdian* (a little). *Nanguo* also co-occurs with various negative feelings, such as *shangxin* (sadness), *tongku* (anguish), *shengqi* (annoyance), *beishang* (sorrow), *weiqu* (grievance), *zhenjing* (shock), *shiwang* (disappointment), *haipa* (fear), *houhui* (regret), *xiukui* (shame) etc. A glance of the concordance of *shangxin* reveals that *houhui* and *xiukui* seldom co-occur with *shangxin*, which might suggest that *nanguo* is the result of the general evaluation of what happened, whether it is caused by oneself or someone else, while *shangxin* is more likely to emphasise the pain on ‘ME’ caused by other people or some external forces.

The last column gives its other collocates which help understand the meaning of *nanguo*. We find *shangxin*, *xiangqi/xiangdao* (think of), *shi/rang/jiao/ling* (cause/make), *ci* (number of times) and *yizhen* (a short period of time) here. Like *shangxin*, the feeling of *nanguo* can be mitigated or overcome after being consoled (*anwei*) or persuaded (*quan*). *Rizi* (life) is a collocate for its literary meaning---‘difficult to live’ (have a hard

time). It is clear from its collocates that *nanguo* is overwhelmingly used in its metaphorical sense in the modern discourse.

Let us come back to the first column of the table, from which it can be seen that, like *shangxin*, *nanguo* is also something one feels (*gandao/ganjue/juede*) in the heart (*xinli/xinzhong*). *Nanguo* is more experienced as a *xinqing* (frame of mind). *Nanguo*, like *shangxin*, is also an emotion involving physical manifestations, as suggested by its collocates: *lianshang/lian* (on the face/face), *yangzi* (appearance), *kandao/kanjian* (see), *ku* (cry), *lei* (tears), *liulei* (shed tears) etc. *Nanguo* appears to be often associated with tears or crying; however, there is a common expression in Chinese:

- 14) *nanguo de dou xiang ku*
 ‘so sad that one wants to cry/weep’

Obviously, *nanguo* is a less intense feeling than *shangxin*. Usually, it is a feeling lasting not very long, but someone may feel *nanguo* for a long time if his or her parent or even a beloved pet died. You may feel *nanguo* because your daughter (suppose you have one) says she is leaving for another country but you want her to stay with you, because your little child is not well-behaved, or because the teacher says your son has no talent for college, because you failed in an exam or an interview, or when you are listening to your friend’s *sad* story etc.

Shangxin and *nanguo* occur in the CCL corpus with almost the same frequency (*shangxin* 4219 hits/ *nanguo* 4152 hits). Table 5.3.4 summarises their collocates indicating features of extroversion and introversion:

Table 5.3. 4 Comparison of the collocates for *shangxin* & *nanguo*

	<i>shangxin</i> 4219 hits	<i>nanguo</i> 4152 hits
--	---------------------------	-------------------------

Intro- version	<i>gandao</i> 143 <i>ganjue</i> 13 <i>juede</i> 60 <i>xinli/xinzhong</i> 17/12	<i>xinli</i> 622 <i>xinzhong</i> 152 <i>xinqing</i> 36 <i>gandao</i> 355 <i>juede</i> 146 <i>ganjue</i> 17
	Total 245 Percentage 58‰	Total 1328 Percentage 320‰
Extro- version	<i>yangzi</i> 30 <i>kan/kandao</i> 87/27 <i>lainshang/lian</i> 16/13 <i>ku/kuqi</i> 436/33 <i>lei/yanlei</i> 75/72 <i>leishui</i> 18 <i>luolei/kiulie</i> 74/27 <i>tongku</i> 35	<i>yangzi</i> 55 <i>lian</i> 15 <i>kandao</i> 108 <i>kanjian</i> 22 <i>ku</i> 69 <i>lei</i> 18 <i>lüelei</i> 12
	Total 943 Percentage 224‰	Total 299 Percentage 72‰

From the numbers and percentages of the extroversion and introversion collocates for *shangxin* and *nanguo* shown in the table, it can be clearly seen that *shangxin* tends to be more shown outwardly than *nanguo* (224>72). *Nanguo* is more likely to be used to describe what is happening internally than *shangxin* (320>58). As mentioned earlier, in Chinese, it is believed that the behaviour of the heart has a lot to do with a person's emotional state. If you feel *nanguo*, your heart will feel uncomfortable. *Shangxin* is a metaphorised word which has been frequently used independently, while *nanguo* is used to describe what is taking place in the heart in a metaphorical way. It is undergoing the change. This is why *nanguo* co-occurs more frequently with *xinli* or *xinzhong* and is less often used as an attributive.

5.2.3.2.3 *Youshang*

The last word I will examine in this section is *youshang*. It is a combination of *you* (worried) and *shang* (sorrowful), meaning 'youchou (worried) beishang (sorrowful)' (XHCBL), so it is also associated with the fatalistic view. 1175 hits of *youshang* are found in the CCL corpus and its significant collocates are shown as follows:

Table 5.3. 5 Significant collocates of *youshang* in the CCL corpus

Introversion	Extroversion	Verb & modifier	Feeling	Others
<i>gandao</i> 44	<i>yanjing</i> 27	<i>chongman</i> 44	<i>tongku</i> 33	<i>si</i> 15
<i>juede</i> 11	<i>lianshang</i> 23	<i>manhuai</i> 9	<i>fannao</i> 12	<i>zhong</i> 43
<i>neixin</i> 12	<i>biaoqing</i> 16	<i>liulu</i> 10	<i>beitong</i> 10	<i>zhezong</i> 18
<i>xinzhong</i> 14	<i>yanshen</i> 14	<i>luchu</i> 8	<i>gudu</i> 9	<i>nazhong</i> 15
<i>xinli</i> 19	<i>shenqing</i> 13	<i>shenshen</i> 12	<i>juewang</i> 9	<i>fen</i> 7
<i>xinling</i> 11		<i>dandande</i> 18	<i>fennu</i> 9	<i>meili</i> 25
<i>xinqing</i> 10		<i>si</i> 9	<i>kongju</i> 9	
<i>qingxu</i> 11		<i>jifen</i> 6	<i>beiai</i> 9	

The table shows that *youshang* is semantically associated with both introversion-words, such as *gandao/juede* (feel), *xin/xinli/xinzhong* (in the heart), *xinling* (mind), *xinqing* (frame of mind), *qingxu* (mood), and extroversion-words, like *yanjing* (eye), *lianshang* (on one's face), *biaoqing/shenqing* (facial expression), *yanshen* (expression in one's eyes). It semantically prefers verbs of 'being brimming with', typically *chongman* (be filled with) and *manhuai* (be filled in one's bosom), and verbs of 'revealing', such as *liulu*, *louchu*. Its typical modifiers include *shenshen* (deep), *dandan de* (faint), *si* (a minute quantity) and *jifen* (small amount). *Si* (lit. silk) and *jifen* (lit. a few + a unit of length) are Chinese specific literary modifiers to describe a slight degree of intensity for emotions.

Like *shangxin* and *nanguo*, *youshang* also tends to co-occur with other negative feelings, such as *tongku* (anguish), *fannao* (vexation), *beitong* (grief), *gudu* (loneliness), *juewang* (despair), *beifen* (grief and indignation), *fennu* (indignation), *kongju* (fear), *beiai* (sorrow). The last column gives *youshang*'s other collocates which are helpful for revealing its meaning. As a word to denote sadness, unavoidably, *youshang* is sometimes associated with *si* (death). One unique feature for *youshang* is that it is more likely to co-occur with classifiers, like *zhong* (*zhezong*/this kind, *nazhong*/that kind)

and *fen* (portion), when it is used as a noun. Consider the following examples from the

CCL:

- 15) 这个“神秘”的音乐组合以他们**优美**(*youmei*) exquisite、**宁静**(*ningjing*) 而带着**忧伤**(*youshang*)的音乐吸引了众多的媒体和乐迷
- 16) 似一幅**温和**(*wenhe*)、**优雅**(*youya*) 而又让人品出一缕淡淡**忧伤**(*youshang*)的风景画
- 17) 擦亮前额使我的**忧伤**(*youshang*)也**楚楚动人**(*chuchudongren*)
- 18) 她显得很**文静**(*wenjing*)，**微笑**(*weixiao*) 的脸上带着几丝淡淡的**忧伤**(*youshang*)
- 19) 她的笑里有一种**宁静**(*ningjing*)、**纯真**(*chunzhen*)与**忧伤**(*youshang*)糅合在一起的东西，使天成感动

All the Chinese words in bold shown above are absolutely positive: *youmei* (exquisite), *ningjing* (calmness), *wenhe* (gentle), *youya* (elegant), *chuchudongren*¹⁷ (moving) and *wenjing* (gentle and quiet), *weixiao* (smiling) and *chunzhen* (innocence). As *youshang* is often associated with concepts with very good connotations, we may argue that it has a less negative semantic prosody and is not perceived as a fully negative emotion. This might be due to its constituent character *you*, which was a quite positive concept in ancient Chinese. We will come back to this issue when we deal with *youyu*.

5.2.3.3 Comparison

The concept of *sadness* has been well researched in the literature. The claim that *sadness* is one of the universal human emotions has also been challenged by many scholars (see Levy, 1973; Stearns, 1993; Wierzbicka, 1999). Their studies show that some languages do not have corresponding words to *sadness*, and others only have some words which are roughly equivalent to it. Wierzbicka claims that people ‘may not be conscious of the reason for the *sadness*’ (Wierzbicka, 1999: 62). Ekman and Friesen

¹⁷ *Chuchudongren* is a descriptive adjective with positive semantic prosody.

(1975: 117) argued that '*sadness* is a passive, not an active feeling', suggesting one is resigned to what happened.

After examining so many sad-feeling words in Chinese, I would argue that none of them is equivalent to *sad*. *Sad* seems to be a generic term covering many sad-related concepts, such as *sorrow*, *grief*, *heartbreak*, *mourn*, *doleful*, *woeful* etc. Each of them differs in many dimensions from others, but the basic semantic component seems to be the same, i.e. sadness. In Chinese, however, there is no such a word which can cover all those sadness-denoting words. Besides, *sad* means 'feeling unhappy', 'showing unhappiness' and 'making you unhappy'. It can be used to describe unhappy feelings, unhappy expressions or things and situations that make people unhappy. Chinese sadness words can only describe people's feelings, expressions and some facial parts or physical events. For example, Chinese sadness words are seldom used to modify a face, such as **nanguo de lian*. In the sense of feeling unhappy, *shangxin* seems to be the closest one since it is often used to explain other sadness words in the dictionary definitions. But, *shangxin* is more intense and extroverted than *sad*. *Nanguo* is similar in intensity to *sad*, but more specific because it focuses on the emotional discomfort brought by what happened. In addition, *nanguo* is less formal. *Youshang* is a Chinese specific concept which is a mixture of worry and sorrow. Like *sad*, someone who feels *youshang* may not be conscious of the reason for *youshang*. *Youshang* is also a passive feeling, i.e. the experiencer of *youshang* may be resigned to what happened. *Youshang* is often associated with *pingjing* (calmness) and *meili* (beauty), and sometimes it is viewed as something people may enjoy. *Youshang* and sadness are similar in being non-

specific and passive. But *youshang* is more intense, more formal, more literary, more introverted than sadness.

5.2.4 *Heartbreak** and their Chinese equivalents

5.2.4.1 *Heartbreak**

Heartbreak is a concept metaphorically conceptualised in terms of a bodily image, namely 'heart is broken'. Let us look at the relevant entries in the dictionaries:

LDCE:

heartbreak: *n* [U] great sadness or disappointment

heartbroken: *adj* extremely sad because of something that has happened

broken-hearted: *adj* extremely sad, especially because someone you love has died or left you

heartbreaking: *adj* making you feel extremely sad or disappointed

OALECD:

break sb's/one's`heart make sb/one feel very sad 使某人/自己/很伤心`

heart-break *n* [C, U] (cause of) very great unhappiness 很大的不幸; 造成很大不幸的原因: *She's had her share of heart-break(s).*她也有她自己的伤心事.

heart-breaking *adj.*

heart-broken *adj* (of a person) feeling great sadness (指人) 极其伤心的

COBUILD:

Heartbreak is very great sadness and emotional suffering, especially after the end of a love affair or close relationship.

Someone who is **heartbroken** is very sad and emotionally upset.

Someone who is **broken-hearted** is very sad and upset because they have had a serious disappointment.

Something that is **heartbreaking** makes you feel extremely sad and upset.

A summary of the above definitions shows that *heartbreak* is great sadness mainly caused by disappointment, the end of a close relationship or someone's death. In OALECD, ironically, it is translated as *buxing*, which means 'misfortune'. In the example, fortunately, its translation *shangxinshi* does not deviate too much from its

real meaning. According to this dictionary, the Chinese equivalent of *heart-broken* is *jiqi* (extremely) *shangxin*.

Now let us turn to the corpus search (*heart-break* 6, *heartbreak* 1157). A close examination of its concordance and its top 50 collocates reveals that *heartbreak* colligates strongly with the definite article *the* (226), possessive adjectives, typically *his* (28), *her* (24), *their* (10), prepositions, predominantly *of* (221), *for* (80), *at* (29), *over* (14). Its word sketch shows that the verbs it often co-occurs include *suffer* (49), *face* (13), *experience* (3), *cause* (7), *bring* (7) etc. *Cause* and *bring* appear to show that there is an explicit reason for *heartbreak*. Consider the following examples:

- 1) To many people, such a recognition requires considerable courage, for casting away anything may seem an intolerable action, just as failing at anything is a great **heartbreak**.
- 2) Mr Justice Waite described his ruling as a happy ending, but warned that people who failed to go through government-approved adoption agencies faced a grave risk of disappointment and even **heartbreak**.
- 3) partnership - a deep relationship with another person is essential to the Libran 's very being , and all too often they rush into total commitment before they are really ready for it; such partnerships can often end in **heartbreak** and disaster .
- 4) Jealousy flamed in her that she had never known that part of his life, that it belonged to his wife, Elizabeth. The blood was pounding in her cheeks. She put up her hands to cover them. ` I 'm just someone ... to pass the time with. I blame myself entirely. Anna thinks I 'm stupid, that it 's pointless , that there 's nothing in it for me but **heartbreak**. 'The pain was unendurable.

Examples 1, 2 and 4 seem to suggest that the principal determinant of *heartbreak* is disappointment, either about oneself or someone else. Example 3 seems to show that the end of a close or deep relationship, whether partnership or love, usually leads to *heartbreak*. One might wish the relationship to continue and it is the disappointment about the person with whom one used to have the relationship that leads to the feeling

of *heartbreak*. However, for the *heartbreak* caused by someone's death, the loss, instead of the disappointment, is the main factor.

Now let us look at *heartbroken*. 1420 hits of *heartbroken** (*heartbroken*, 1197; *heart-broken*, 43; *brokenhearted*, 56; *broken-hearted* 124) are found in the BOE. A close inspection of its concordance and its top 50 collocates reveals that *be* (434) dominates the picture. It also has a colligation with prepositions, typically *over* (31), *at* (38). It is interesting to find *parent* (34), *family* (30), *mum* (25), *mother* (21), *father* (16), *dad* (14), *wife* (10), *relative* (6), *husband* (6), *brother* (5) in the top collocates. It suggests that *heartbroken** has a semantic preference for family connections because the 'heartbroken' feeling is closely associated with family members. Both *feel* (7) and *look* (7) are found to occur with *heartbroken**, suggesting that *heartbroken* is both extroverted and introverted. 'Be left heartbroken' is a recurrent pattern which occurs 40 times in the BOE. It shows that getting *heartbroken* is the immediate result of something bad. Its word sketch shows that *heartbroken* is often modified by maximisers, such as *absolutely* (21), *utterly* (3), *truly* (3) and *totally* (3), showing that *heartbroken* can only be reinforced but not weakened. We will come back to this issue when we talk about its Chinese equivalents.

For *heartbreaking*, 1446 occurrences (*heartbreaking*1214+*heart-breaking*232) are found in the BOE. An inspection of its concordance listing and its top 50 collocates shows that it has colligations with *be* (369), the article-determiners *a* (239), *the* (158), possessive adjectives, typically *his* (20), *their* (17), and the preposition *for* (66). Its word sketch shows that it is often modified by *absolutely* (5), *truly* (3), *especially* (4),

pretty (4), *particularly* (3). It should be noted that, different from *heartbroken*, *heartbreaking* can be modified by the compromiser *pretty*, suggesting that *heartbreaking* is a less strong word than *heartbroken*. According to its word sketch, *heartbreaking* is semantically associated with negative words that can make one extremely sad or upset, such as *loss* (51), *defeat* (28), or words that have such a quality, like *story* (21), *decision* (19), *letter* (17), *news* (13), *moment* (12), *performance* (10), *plea* (9) etc.

Only 81 instances of *heartbreakingly* are found in the corpus. Although it does co-occur with adjectives with negative connotations, such as *naff* (1), *ineffective* (1), *fragile* (1), *inadequate* (1), *awful* (1), *stupid* (1), *sad* (1), *thin* (1), it co-occurs more frequently with positive words, like *beautiful* (16), *lovely* (3), *handsome* (2), *delicious* (1), *unspoilt* (1), *touching* (1), *hopeful* (1), *brave* (1), *pleasant* (1), *poetic* (1). So it may be argued that *heartbreakingly* has evolved into a pure intensifier with a more positive semantic prosody.

5.2.4.2 *Xinsui*

30 hits of *heartbreak** (incl. *heartbreaking*, *broken(-)hearted*, *heartbroken*) are found in the BPC. Below is the result of my search:

Table 5.4. 1 Chinese equivalents of *heartbreak** in the BPC

Chinese equivalents	Frequency
<i>shangxin</i> :	20 (<i>shang tou le xin</i> 1+ <i>shangxin tou le</i> 1)
<i>xinsui</i>	3
<i>beitong</i>	3
others	4
Total	30

It is understandable that the great majority of the instances of *heartbreak** are translated as *shangxin* because it is almost a cover term to indicate a sad emotional state. If the translator can not find an appropriate word to describe a sad state, s/he would resort to this word and put an intensifier before it. For example:

- 5) ‘一想起你这么快便忘记了自己的教养，我就**伤心透了** (*shangxin tou le*)’。
‘I am **heartbroken** to think that you could so soon forget your rearing’.
- 6) 并且波琳上星期还写了信来，说他名声很坏，在查尔斯顿，连他自己家里也没有接待他，只是他那位**伤透了心** (*shang tou le xin*) 的母亲例外。 ...and Pauline wrote me only last week that he is a man of bad repute and not even received by his own family in Charleston, except of course by his **heartbroken** mother.

In the above examples, (*shangxin*) *tou le*/*(shang) tou le (xin)* are used to indicate the degree of *shangxin*. *Tou* literally means ‘pass through’. The whole construction means the heart is penetrated through, i.e. the heart is badly hurt. In some contexts, if someone you love has died, you will be in *heartbreak*. In the following examples, *heartbreak** is translated as *beitong* (3), which is a word usually denoting grief due to the death of a beloved person:

- 7) 她虽然处于**悲痛**(*beitong*)之中，然而一想到和他见面，而她怀的又是另外一个男人的孩子，就感到不寒而栗。 Even in her **heartbreak**, she shrank from the thought of facing him when she was carrying another man's child.
- 8) 她本来会奉献给教堂的那分**悲痛**(*beitong*)和无私，如今都全部用来服务于自己的儿女和家庭以及那位带她离开萨凡纳的男人了... The **heartbreak** and selflessness that she would have dedicated to the Church were devoted instead to the service of her child, her household and the man who had taken her out of Savannah ...
- 9) 如今他那双光亮的马靴踏上了苦难的道路，那儿充满了饥饿、疲惫、行军、苦战、创伤、**悲痛**(*beitong*)等等，像无数狂叫的恶狼在等着他，最后的结局就是死亡呢。 Now he had set his varnished boots upon a bitter road where hunger tramped with tireless stride and wounds and weariness and **heartbreak** ran like yelping wolves. And the end of the road was death. He need not have gone.

Another Chinese near-equivalent of *heartbreak** is *xinsui* (3), which literally means

‘heart is broken into pieces’. Look at the examples from the BPC:

- 10) 她呼唤着孩子的名字，有时头抬起来整整有一分钟时间那么长听着，然后无力地呻吟着一头倒在床上。见此情形，大家都说真叫人心碎(*xinsui*)。People said it was *heartbreaking* to hear her call her child, and raise her head and listen a whole minute at a time, then lay it wearily down again with a moan.
- 11) 人们从媚兰那年轻的脸上可以看出，她对过去的一切是忠贞不渝的。这使人们会暂时忘记自己一伙人中那些使人愤怒、害怕、心碎(*xinsui*)的败类。When they looked into her young face and saw there the inflexible loyalty to the old days, they could forget, for a moment, the traitors within their own class who were causing fury, fear and *heartbreak*. And there were many such.
- 12) 我知道，她在爱你时，看到了跟她同龄的母亲，也在爱着她，看到了跟我同龄时的你，也在爱着我。她爱她心碎(*xinsui*)的母亲，她爱那经历了可怕的考验和成功的恢复过程的你。I know that in loving you she sees and loves her mother at her own age, sees and loves you at my age, loves her mother *broken-hearted*, loves you through your dreadful trial and in your blessed restoration.

Xinsui is not included in XHCBL, probably because it is not a frequent word. Only 430

hits of *xinsui* are found in the CCL corpus and its significant collocates are shown

below:

Table 5.4. 2 Significant collocates of *xinsui* in the CCL corpus

Collocate	Frequency
<i>ling(ren)163/rang 36/shi 50/jiao20</i>	269
<i>si 14</i>	14
<i>ku/kuqi</i>	21
<i>gushi</i>	6
<i>xiaoxi</i>	7
<i>muqin</i>	9
<i>qiliang</i>	5
<i>chang</i>	14
<i>lie</i>	6

It is interesting to find the most significant collocates for *xinsui* are *ling/rang/shi/jiao*

(cause/make). It can be seen that in more than half of the occurrences (269/430, 63%),

xinsui is used in a causative construction, whose function in English can be fulfilled by

a single adjective, such as *sad*, *sorrowful*, *heartbreaking*. Of course, some other expressions can also be translated as *xinsui*. For example:

- 13) 我的天哪！心碎(*xinsui*)的人多了，也没见谁去当修女。就拿我来说吧，我送掉了一个丈夫’。 ‘Lots of people's *hearts* have been *broken* and they didn't run off to convents. Look at me. I lost a husband.’
- 14) 你认为我可以丢下媚兰和孩子自己跑掉，就算我恨他们两个人，难道我能让媚兰心碎(*xinsui*)？...’ ‘Do you think I could go off and leave Melanie and the baby, even if I hated them both? *Break* Melanie's *heart*? ...’
- 15) 在路上看见她路过，把消息带回家里，我爸爸便心碎(*xinsui*)而死。他满腹冤屈，却一个字也没说。 I saw her pass me on the road. When I took the tidings home, our father's *heart burst*. He never spoke one of the words that filled it.

It seems that in the English original texts if expressions like ‘hearts have been broken or break someone’s heart’ are used to describe the great sadness, then probably, their Chinese translation would be *xinsui*. If *heart burst* is used instead of *heart break*, then *xinsui* is also highly likely to be used to translate it. The reason for this would be that cleft expressions are more intense than single words.

In Chinese, *xinsui* is also closely related to *si* (death) and *ku/kuqi* (cry/weep). *Xinsui* is often used in a causative construction to describe things that make people extremely sad, such as *lingren xinsui de gushi* (sad story) or *lingren xinsui de xiaoxi* (sad news). *Muqin* (mother) often co-occurs with *xinsui* probably because a mother is often sentimental and tends to be the experiencer of *xinsui*. A *qiliang* (desolate) place may make one *xinsui* because it is unbearable to see that exceedingly sad scene and one’s feeling in such a situation is also conceptualised as *xinsui*. *Chang* (intestine) and *lie* (crack) appear in the table because in the Chinese culture, *xinsui*, *chang duan* (intestine-broken) and *dan lie* (gall bladder-crack) are common metaphorical images to

understand great sadness. One thing that needs to be noted is that *xinsui* is not always used to denote a sad feeling, sometimes it is positive. Consider the following examples:

- 16) 一片令人心碎(*xinsui*)的温柔
'*lingren xinsui de gentleness*'
- 17) 那声声妈妈，叫得好甜哟，叫得让人心碎(*xinsui*)哟...
'the calling of 'mama' was so sweet and *lingren xinsui*'
- 18) 美得令人心碎(*xinsui*)
'so beautiful that it made someone *xinsui*'

The above examples show that *xinsui* is not necessarily bad or sad. It is used to describe an immensely strong feeling, regardless of whether it is positive or negative.

5.2.4.3 Comparison

Heartbreak is associated with great sadness, caused by breaking up, the ending of a close relationship, the death of a beloved person, or something very bad that happened.

Heartbreak is absolutely personal, i.e. it is characterised by 'something bad happened TO ME'. It is more an internal state, a mental suffering or pain. It is more related to family or a close relationship. Its Chinese counterparts could be *shangxin*, *beitong* or *xinsui*. Among them, *shangxin*, often with its premodifier, is the most common translation equivalent in Chinese. *Shangxin* alone is not intense enough to express *heartbreak*. *Beitong* is only used when it is associated with someone's death. *Xinsui* is an intense feeling mainly used in causative constructions like 'lingren xinsui de'. In addition, *xinsui* is not necessarily sad or bad. Therefore, in Chinese, there is no word that corresponds to *heartbreak* or any of its cognate. English and Chinese conceptualise sadness by associating the heart with different metaphorical images. *Sui* means '完整的

东西破成零片零块 break to pieces; smash' (XHCBL). Let us look at the definition of

break in LDCE:

If you break something, you make it separate into two or more pieces, for example by hitting it, dropping it, or bending it.

It is clear that 'two or more pieces' is used to describe the extent to which something is broken. So in terms of intensity, *heartbreak* is half way between *shangxin* (heart-hurt) and *xinsui* (heart-fragmented). Usually, the translator increases the intensity of *shangxin* by adding an amplifier to it to match *heartbreak*. Sometimes, *xinsui* is used instead, but *xinsui* is far less frequent than *shangxin*, and it has a literary flavour. To sum up, Chinese and English encode different meanings in their sadness words, which are not corresponding to each other at all. We have to rely on the corpus data to find out their subtle nuances of meaning.

5.2.5 *Doleful** & *woe** and their Chinese equivalents

5.2.5.1 *Doleful**

The dictionary entries for the *doleful* family are shown below:

LDCE:

doleful *adj formal* very sad [Date: 1200-1300; Origin: *dole* 'sadness' (13-19 centuries), from *Latin* *dolere* 'to be sad'] ■ *a doleful song about lost love*

OALECD:

doleful *adj* sad; mournful 悲伤的; 令人沮丧的: *a doleful face, manner, expression, etc* 愁眉苦脸/垂头丧气的样子/忧郁的表情.

COBUILD:

doleful [ADJ]

A ***doleful*** expression, manner, or voice is depressing and miserable. = mournful • ***dolefully*** [ADV] ADV with v

LDCE says that *doleful* derives from *dole*, which comes from a Latin word *dolere*, meaning 'to be sad'. According to the other two dictionaries, *doleful* is very close to

mournful. 157 hits of *doleful* are found in the BOE. According to its top collocates¹⁸, its colligational properties include the article-determiners *a* (34), *the* (23), possessive adjectives, typically *his* (7), *their* (5), *our* (5), and nouns. *Doleful* semantically prefers words relating to external, especially visible, manifestations, such as *expression* (8), *eye* (6), *face* (4) and *look* (4). It implies that *doleful* describes outward expressions. *Be* or any of its variant form, like *is*, *are*, *am*, does not occur before *doleful*, suggesting that *doleful* is usually used as an adjectival modifier. Unlike *sad*, it is seldom used for predicating.

Now let us turn to its adverbial form. *Dolefully* only occurs 88 times in the BOE. Its top 50 collocates show that *dolefully* is semantically associated with physical action verbs, like *say* (11), *nod* (3), *stare* (3) etc. An edited concordance listing is shown below:

- 1) to the bar. Jewel shook her head dolefully. <p> `Me neither. She better
- 2) Sympathizing, Mabel shook her head dolefully. `Just sittin' there, huh?'
- 3) e Earthmen.' Sefan shook his head dolefully. `This is a bad business.'
- 4) them under fifty, shook their heads dolefully in perfect agreement with thei
- 5) race ended. Keeler **glanced** back dolefully and shook his head. Suckers .
- 6) see you, others laze in the sun and dolefully **watch** you wander by. These are
- 7) you like it, but.. 'She **looks** around dolefully, a smile for the cat, '...why'

A simple glimpse of the above lines shows that 'shook one's head *dolefully*' is a frequent pattern. 'Keeler glanced back *dolefully*' in line 5, 'others laze in the sun and *dolefully* watch you wander by' in line 6 and 'looks around *dolefully*' seem to suggest that *dolefully* can collocate with any looking verb, such as *look*, *watch*, *glance*.

To determine the Chinese equivalents of *doleful* and *dolefully*, we still need to look at the BPC, in which only 6 instances are found:

¹⁸ The Bank of English gives top collocates based on both raw frequency and t-score. Collocates with low frequency may lead to negative t-scores and are therefore excluded from the list.

- 1) 贾维斯罗瑞因为心绪不宁，也因为他年轻的同伴越来越激动，也曾两次停下步来休息，每次都在一道凄凉(*qiliang*)的栅栏旁边。还没有完全败坏，却已失去动力的新鲜空气似乎在从那栅栏逃逸，而一切败坏了的带病的潮气则似乎从那里扑了进来。Yielding to his own disturbance of mind, and to his young companion's agitation, which became greater every instant, Mr. Jarvis Lorry twice stopped to rest. Each of these stoppages was made at a *doleful* grating, by which any languishing good airs that were left uncorrupted seemed to escape, and all spoilt and sickly vapours seemed to crawl in.
- 2) ‘我，不幸的医生 亚历山大·曼内特，波维市人，后居巴黎，于一七六七年 最后一个月在 巴士底狱凄凉(*qiliang*)的牢房里写下这份悲惨的记录’。’I, Alexandre Manette, unfortunate physician, native of Beauvais, and afterwards resident in Paris, write this melancholy paper in my *doleful* cell in the Bastille, during the last month of the year, 1767’.
- 3) 他为她们又是着急又是痛苦(*tongku*)，日子过得极其缓慢沉重。A disturbed and *doleful* mind he brought to bear upon them, and slowly and heavily the day lagged on with him.
- 4) ‘我的小女儿。’派逊斯答道，神情有些悲哀(*beiai*)，但又自豪。It was my little daughter,’ said Parsons with a sort of *doleful* pride.
- 5) 一棵肃穆的老树对另一棵树悲声(*beisheng*)低吟，仿佛在倾诉树下坐着的这一对儿子的伤心的故事，或是在不得不预告那行将到来的邪恶。...while one solemn old tree groaned *dolefully* to another, as if telling the sad story of the pair that sat beneath, or constrained to forebode evil to come.
- 6) 我觉得，对于早晨伤亡惨重的这帮家伙来说，这支歌是再(*zai*)合适不过(*buguo*)了。And I thought it was a ditty rather too *dolefully* appropriate for a company that had met such cruel losses in the morning.

In the last instance, *dolefully* is used as a degree adverb to modify the adjective *appropriate*. So we will focus on the others. *Dolefully* can be translated as *shangxin (de)* or *beishang (de)*, but usually, in fiction, it is translated as something more literary, like *beisheng diyin* in Example (5). In the first four of them, *doleful* is used as an attributive to modify things: *grating*, *cell*, *mind* and *pride* respectively. Among them, the first two are translated as *qiliang*, which in Chinese is only used to describe surroundings or things, meaning ‘being lonely and can arouse sad feelings’. *Doleful* in *doleful mind* is translated as *tongku* and in *doleful pride* as *beiai*.

Unlike *sad*, *doleful* is hardly used to modify someone who feels sad and only used to modify things that show *sadness* or make people *sad*, such as *look, eyes, face, facial expressions, song* as revealed in the BOE, and *grating, cell, mind, pride* as shown in the above examples. OALECD, as presented earlier, gives 5 different equivalents: *beishang de* (悲伤的), *lingren jusang de* (令人沮丧的) in the definition, and *choumeikulian* (愁眉苦脸), *chuitousangqi* (垂头丧气) and *youyu* (忧郁) in the translations of its examples--- ‘*a doleful face, manner, expression*’. We can see that for a *doleful* face, we find ‘*choumeikulian*’; for a *doleful* manner/expression, *chuitousangqi* and *youyu* are used respectively. These words are chosen to translate these phrases just because they are frequent collocations of the modified words in Chinese. In effect, they all denote quite specific feelings and none of them is equivalent to *doleful*. *Lingren jusang de* is used to match *doleful* because both are extroverted. *Beishang* is used because they are similar in intensity, but the difference between them is that *beishang* is more introverted and *doleful* is more extroverted. *Doleful* has a much broader range of collocations, such as *cell, pride* and the like, but any above-mentioned Chinese word tends to modify a rather limited range of things. So no Chinese word corresponds well to *doleful*. Chinese emotion words are relatively limited in their collocating behaviour, while their English counterparts tend to be more flexible in collocations.

5.2.5.2 *Woe**

This section will look at *woeful, woefully* and *woebegone*. Before examining them, I will first look at *woe*. Below are the entries of *woe* from the three dictionaries:

LDCE: *n* [Language: Old English; Origin: wa]

▣▶**woes** [plural] *formal* the problems and troubles affecting someone

▣▶[U] *literary* great sadness

OALECD: *n* (*dated or fml or joc* 旧或文或谑)

▣▶[U] great sorrow or distress 悲哀; 悲痛; 苦恼 ▣▶**woes** [pl] things that cause sorrow or distress; troubles or misfortunes 引起悲哀或苦恼的事物; 麻烦事; 不幸的事

COBUILD:

1 [N-UNCOUNT] ≠ joy

Woe is very great sadness. (LITERARY)

2 [N-PLURAL] usu with poss

You can refer to someone's problems as their **woes**. (WRITTEN)

These dictionaries provide the same senses about *woe*: ‘great sadness’ (U) and ‘problems’ (PL). Now let us look at what it is like in the corpus. Since only its singular form denotes sadness, I will focus on this form. I got 1092 hits of *woe* (U) from the BOE. A glance of its top 50 collocates reveals that *woe* colligates strongly with possessive markers, such as *of* (260), *'s* (32) (indicating possessive case). Its concordance shows that, in modern discourse, *woe* is usually used in (half-)fixed expressions like *woe betide somebody*, *woe of tale(s)*, *woe is me*, *full of woe*, *in weal and/or woe*, and *add woe* etc.

Now let us turn to *woeful* and *woefully*. About them, the dictionaries say:

LDCE:

woeful *adj*

▣▶very bad or serious = deplorable ▣▶*literary* very sad = pathetic

>**woefully** *adv*

OALECD:

woeful *adj* (*fml* 文)

▣▶full of woe; sad 悲哀的; 伤心的 ▣▶[usu attrib 通常作定语]

undesirable or regrettable; very bad 不合意的; 令人惋惜的; 糟糕的

> **woefully** *adv*

COBUILD:

1 [ADJ] If someone or something is **woeful**, they are very sad.

● **woefully** [ADV] ADV with v

2 [ADJ] usu ADJ n You can use **woeful** to emphasise that something is

very bad or undesirable.

- **woefully** [ADV] usu ADV adj, also ADV before v

From the definitions above, it is quite obvious that *woeful* is basically defined as ‘sad’ or ‘bad/undesirable’, and that *woefully* can be used to modify a verb to indicate a sad manner, or used as a degree modifier. *Woeful* occurs 912 times in the BOE. By inspection of its top 50 collocates, it is clear that *woeful* has colligations with *be* (126), the article-determiners *a* (171), *the* (140), possessive adjectives, typically *their* (37), *his* (21), and determiners, predominantly *this* (11). According to its word sketch, its intensifiers are *truly* (7), *utterly* (2) and *pretty* (4). The words modified by *woeful* include *display* (27), *record* (27), *performance* (23), *start* (13), *lack* (13), *defeat* (6), *showing* (4), *miscalculation* (4), *inadequacy* (3), *ignorance* (3), *calamity* (2), *inconsistency* (2) etc. It can be seen that almost all of these words are associated with its sense of ‘bad/undesirable’. I can hardly recognise any word that is surely related to ‘sad’. Many of them talk about football or other sports. This might be due to the composition of the BOE, which contains a lot of newspaper articles. It shows that, in modern discourse, *woeful* is much more used in the sense of ‘bad’ than ‘sad’.

Now let us move on to the discussion of its adverbial form. 698 hits of *woefully* are found in the corpus. In its top 50 collocates, I find *look* (8) and *head* (6, from *shake/nod one’s head*), which show that *woefully* is used to modify an action verb to indicate an emotional manner. However, its word sketch reveals that, in most of the cases, *woefully* is used to modify an adjective to indicate a degree. The modified adjectives include *inadequate* (162), *short* (74), *underprepared* (22), *ignorant* (14), *ill-unprepared* (7), *deficient* (7), *underpaid* (6), *insufficient* (4), *understaffed* (4), *ill-*

informed (3), undermanned (3), inept (3), outdated (3), misguided (3), inexperienced (3) etc. It is clear from the above list that *woefully* has evolved into an intensifier with a negative semantic prosody.

Now let us turn to look at their Chinese equivalents. *Woe** occurs only 23 times in the BPC. Below is the searching result:

Table 5.5. 1 Chinese equivalents of *woe** in the BPC

Sense	Chinese equivalents
Sad	<i>tongku 6 aiyuan 1, shangxinshi 1, zhuxinqixue 1, aishang 1, beishang 1, nanshou 1, beitung 1</i>
Bad/undesirable	<i>daomei 2, kelian 1, zainan 1, muburendu 1 (in woefully visible)</i>
<i>Chou</i>	<i>beiku 1, kunao 1, choumeikulian 1, choukubukan 1, chourongmanmian 1</i>

As mentioned in Chapter IV, the parallel corpus used for this research is composed of fiction alone. In fictional writing, it is one of the main features to describe people's feelings or emotional states. So, unlike the BOE, no instance of *woefully* from this corpus is related to the degree adverb sense. From the above table, it can be seen that *woe** is often translated as *tongku* in Chinese. The translators use many Chinese words to express the concept of *woe*, but basically, these words can be categorised into three classes: one is connected to the feeling of sadness, like *tongku, aiyuan, aishang, beitung*; the second one is related to something that is bad or undesirable, like *daomei, kelian, zainan*; and the third one is associated with *chou*, which is close to *woes*. Since the focus of this research is to examine the concept of sadness, I will only look at the first group.

Woe is a literary word, so Chinese translators use formal or literary words, like *aiyuan, aishang, zhuxinqixue, beishang, beitung*, to express this concept. *Tongku*

focuses on the unacceptability of the bad events and the sufferer's incompetence to solve the problem. The pain comes mainly from the helplessness of the experiencer. We will discuss this concept in detail in a later section. Translators use *tongku* to express *woe* probably because they view *woe* as a concept between *tongku* and *beishang*. Look at the following example from the BPC:

- 7) Cried he, with a voice that roes over them, high, solemn, and majestic, yet had always a tremor through it, and sometimes a shriek, struggling up out of a fathomless depth of remorse and *woe* -- 'ye, that have loved me!' 他的声音高昂、庄严而雄浑,一直越过他们的头顶,但那声音是从痛苦(*tongku*)与悔恨的无底深渊中挣扎出来的,'你们这些热爱我的人!'

In Chinese, *beishang* is something you suffer but you do not fight against it actively. You just let it turn weaker gradually. But it is quite common to say you should fight against *tongku* and try to get out of it as soon as possible. In the original, 'out of' might be the reason why the translator uses *tongku* instead of *beishang*. Consider another example:

- 8) With a hand's breadth flight, it would have fallen into the water, and have given, the little brook another *woe* to carry onward, besides the unintelligible tale, which it still murmuring about. 只消再飞这几指宽的距离,红字就会落进水里,那样的话,小溪除去喃喃不断诉说着的莫测的故事之外,又要载着另一段哀怨(*aiyuan*)流淌了。

In the above citation, *aiyuan* is used presumably because *aiyuan*, meaning 'aggrieved or resentful', is more literary than either *tongku* or *beishang*.

Let us move on to look at *woebegone*, a compound made up of *woe* and *begone*.

Its dictionary entries are shown below:

LDCE: *adj literary* [Date: 1200-1300; Origin: *woe* + *begone* 'affected' (from *bego* 'to surround, affect' (11-16 centuries), from *Old English* *began*, from *gan* 'to go')] looking very sad ■her *woebegone* expression
 OALECD:

*adj (fml 文) looking unhappy 显出悲伤的; 忧愁的; 愁眉苦脸的: a **woebegone** child, expression, face 悲哀的孩子、表情、面孔.*

COBUILD: [ADJ]

Someone who is **woebegone** is very sad. (WRITTEN)

*She sniffed and looked **woebegone**.*

= miserable ≠ joyful

Both LDCE and OALECD agree that *woebegone* is an emotion characterised by extroversion. COBUILD does not say so explicitly, but its example betrays this feature.

Although its etymology in LDCE shows that it originally meant ‘be affected/surrounded by woe’ (*woe+begone*), its contemporary dictionary definitions all point to its feature of extroversion. The OED Online gives its first sense as follows:

‘Beset with woe’; oppressed with misfortune, distress, sorrow, or grief.

Obs. or arch.

According to the OED, it appears that the above sense is its original meaning, but this sense has become inactive in modern usage. Its corpus data provide further evidence for this change. 71 hits of *woebegone* are found in the BOE. *Expression* (5) and *look* (5), but not *be*, are found its most its top collocates, suggesting that *woebegone* has become an emotion word which is specifically used to describe the exhibition of sadness.

6 hits of *woebegone* are found in the BPC: *choumeikulian*, *chourongmanmian*, *manlianbeishang*, *choukubukan*, *ku’nao*, *keliande*. *Choumeikulian* (sadness+eyebrow+bitter+face) is the Chinese equivalent used in the OALECD and is also the closest in meaning. *Chourongmanmian* (face covered with sadness) is much less frequent and only used in certain contexts. *Manlianbeishang* (full face +sadness) is hardly a word. Both *choukubukan* (sadness+bitterness+unbearable) and *ku’nao* (vexation) more describe the inner emotional state. *Keliande* (poor) can not be taken as its equivalent. So the following discussion will focus on this *choumeikulian*, which occurs 362 times in

the CCL corpus. An inspection of its collocates reveals that it co-occurs predominantly with *shuo* (say/46), *kan/kandao/kanjian/jian* (see: 23/13/6/5), *yangzi* (appearance/21), *biaoqing* (expression/5), all of which indicate that it is closely associated with external manifestations of sadness. It also collocates with *zhengtian/zhengri* (all the day: 19+5) and *zongshi/yizhi* (always: 7+5), both showing that it is a long-lasting emotional state. It co-occurs with *aishengtanqi* (have deep signs) 9 times in the corpus, describing vividly what a sad person looks like and suggesting that it is an emotion of resignation. *Choumeikulian* usually has a cause (*yinwei/because*: 8) and a mother (7) more tends to be in such an emotional state.

5.2.5.3 Comparison

As claimed earlier, there is no Chinese word that corresponds to *doleful*. It is also true of *woeful*. In English, *sad* can be used to modify a person who is sad, an expression, facial part or anything showing sadness, or a thing causing a sad feeling, but *doleful* is mainly used to modify a thing that shows sadness, like *doleful song*, *doleful eyes*. Sometimes it is also used to modify a thing that makes people sad, like *doleful cell*, but it is hardly used to describe a person who feels sad, such as *a doleful boy*. *Woeful* may be used to talk about a person, but not often. Both *dolefully* and *woefully* can be used to describe an action to indicate a manner, such as *say*, *groan*, *talk*, *nod*, *shake one's head* etc. Both *doleful* and *woeful* are very formal and literary words. *Doleful* is used to modify things causing or showing sadness and Chinese has many words to describe such things, so its Chinese equivalents are not restricted to a few words, but none of them corresponds

exactly with it. In the case of *woe*, it is even more complex. From the previous discussion, it seems that *woe* is a concept which combines *tongku* and *beishang*, so it makes the matching job more challenging. *Woebegone* is a literary word which is specifically used to describe the external manifestation of a sad person. Chinese tends to combine emotion words and facial organs to give an account of what a person looks like when s/he is very sad, such as *manlianbeishang*, *chourongmanmian* and *choumeikulian*, of which *choumeikulian* has developed into a word due to its high frequency.

It should be noted that the parallel corpus I used for this research is not big, so it only yields a small number of instances of *woe** and *doleful**. The claims based on this corpus still await a further test from bigger parallel corpora. What we can surely see is that English and Chinese encode quite different meanings in these words, which makes it impossible to match the two languages on a word-to-word basis. If a word is looked at in isolation, it shows a complicated picture. However, if it is put in a concrete context, its ambiguity will disappear. In translating, the translator can only decide which aspect is emphasised by relying on the context and select the closest word to it.

5.2.6 *Mourn** and their Chinese equivalents

5.2.6.1 *Mourn**

In this section, we will discuss the *mourn* family. Before we examine *mourn* in the corpus, let us look at what the dictionaries say about it:

LDCE:

☞ to feel very sad and to miss someone after they have died = **grieve**
for ☞ to feel very sad because something no longer exists or is no longer as good as it used to be

OALECD:

v [I, Ipr, Tn] ~ (**for/over sb/sth**) feel or show sorrow or regret for the loss of sb/sth 因丧失某人/某事物/而悲痛或表示哀悼

COBUILD:

1 [VERB] V n, V for n, V

If you **mourn** someone who has died or **mourn for** them, you are very sad that they have died and show your sorrow in the way that you behave.

2 [VERB] V n, V for n

If you **mourn** something or **mourn for** it, you regret that you no longer have it and show your regret in the way that you behave.

From LDCE's definitions, i.e. 'to feel very sad and to miss someone after they have died' and 'to feel very sad because something no longer exists or is no longer as good as it used to be', it can be seen that, the determinant of '*mourn*' is loss. The loss may be material or mental. It may refer to the loss of something or somebody. It seems that the person who mourns feels sad because s/he is always thinking about the past when what is mourned was still there and comparing the difference. OALD includes the meaning of 'show one's sorrow or regret', which is consistent with COBUILD's explanation: 'If you mourn someone who has died or mourn for them, you are very sad that they have died and show your sorrow in the way that you behave'. So, it can be said that *mourn* means two layers of meaning: one is that you are sad and the other is that you show your sadness. Now Let us turn to the BOE. A search for *mourn* generates many instances of *mourning* which are tagged as VVG but are actually nouns, as exemplified as below:

brbooks ' In France, Napoleon Bonaparte ordered ten days *of mourning* .

brbooks ' She touched Sheila's arm across the table, her long slender arm in its black stretch lycra, the colour *of mourning* , clinging to life.

brbooks No sense *of mourning*.

brbooks The Archbishop of Canterbury, the Catholic Bishop of Westminster and the Chief Rabbi issued a joint statement calling for a national day *of mourning* and prayer; flags on all public buildings were flown at half-mast; and a Book of Remembrance opened at the American embassy in Grosvenor; by the end of the day it had been signed by three thousand people.

brbooks `That period *of mourning* was a short one, mate.

brbooks A separation was a bereavement and called for a period *of mourning*, she explained to one of her girlfriends who accused her of being `unsociable' when she declined an invitation to a billionaire's private Caribbean island for a week's vacation.

The discussion of the inaccuracy of its POS tagging is out of the scope of the research.

Since it is inaccurate to examine *mourn* as a lemma, I search for *mourn/mourns/mourned* (2561 hits) in the corpus. An examination of their top 50 collocates shows that *mourn* has colligations with the prepositions *for* (159) and *over* (16), possessive adjectives, typically *his* (91), *her* (85), *their* (71), and pronouns, predominantly *her* (85), *she* (50), *him* (48), *he* (42). The nouns associated with *mourn* include *nation* (52), *family* (42), *death* (36), *world* (36), *loss* (30), *crowd* (19) and *victim* (30). As expected, the most significant component of *mourn* is *loss*. It is often related to someone's death. *Mourning* could be personal. You *mourn* them because they are very close to you. After losing them, you miss them and feel sad. However, we can also find *victim*, *nation*, *world* etc., which show that *mourning* could also be social, communal or culturally-mediated. In other words, people *mourn* not only for those they love deeply, but also for those about whom they think they should show their sorrow.

Its word sketch shows that *mourn* often co-occurs with *privately* (11), *openly* (5), *publicly* (5), *together* (11), *alone* (5), *deeply* (7). So *mourning* can happen in a private or public way; people *mourn* together or alone. The only intensifier is *deeply*.

Now we come to the discussion of *mourning*:

LDCE: n [U]

▣ great sadness because someone has died
in *mourning* (=feeling great sadness)

▣ black clothes worn to show that you are very sad that someone has died

OALECD:

n [U] black or dark clothes worn as a (conventional) sign of grief at sb's death 丧服 (黑色或深色的) :

COBUILD:

1 [N-UNCOUNT]

Mourning is behaviour in which you show sadness about a person's death.

2 [PHRASE] usu v-link PHR

If you are in **mourning**, you are dressed or behaving in a particular way because someone you love or respect has died.

LDCE says *mourning* means both 'great sadness because some has died' and 'black clothes worn to show that you are very sad that someone has died'. OALECD only gives its meaning as clothes. COBUILD says '*mourning* is behaviour in which you show sadness about a person's death'. These three dictionaries differ in the explanation of 'mourning' and the phrase 'in mourning'. Let us look at what it is like in the corpus. 4038 hits of *mourning*, either tagged as VVG or NOUN, are found in the BOE. A careful inspection of its top 50 collocates reveals that *mourning* has colligations with prepositions, typically *of* (854), *in* (546), *for* (418), *over* (35), the definite article *the* (520), and possessive adjectives, predominantly *her* (70), *his* (40), *their* (37). The adjectives associated with *mourning* include *private* (13), *public* (33), *official* (54), *national* (118) and *deep* (29). Its nominal collocates contain *family* (17), *period* (102), *process* (24), *clothes* (21), *ritual* (24), *ring* (11). It seems that *mourning* in the BOE is more talked about as a public *mourning* and this is due to its composition. Overall it has a semantic preference for words related to the *mourning* event.

Next let us turn to look at *mournful* and *mournfully*:

LDCE:

mournful: very sad

OALECD:

mournful *adj* (often derog) sad; sorrowful 悲哀的; 令人悲痛的:

COBUILD:

1 [ADJ] If you are **mournful**, you are very sad.

● **mournfully** [ADV] usu ADV with v

2 [ADJ] A **mournful** sound seems very sad.

594 hits of *mournful* are found in the BOE. Its top 50 collocates show that *mournful* colligates with the article-determiners *a* (122), *the* (109), and nouns, typically *sound* (22), *tune* (14), *cry* (13), *look* (11), *wail* (10), *expressions* (10), *face* (9), *tone*(9), *song* (9), *eye* (8), *voice* (7), *dirge* (6), *ballad* (6), *melody* (5), *music* (5), *howl*(5), *note* (5), *moan* (5) etc. From its collocates, we can see that it has a semantic preference for words which are related to facial expressions, like *face*, *eyes*, *look*, *expression*, or sound, such as *sound*, *tune*, *cry*, *wail*, *tone*, *song*, *voice*, *dirge*, *ballad*, *melody*, *music*, *howl*, *note*, *moan*. Its word sketch shows that the verbs collocating with *mournful* are mainly *look* (11) and *sound* (2), and the only degree modifier is the compromiser *rather* (8). So it may be argued that *mournful* is mainly external because it has visible or audible features to show the sadness, instead of only suffering it mentally; in other words, it is an externalised emotion, which is mainly expressed by facial expressions, sound or music. ‘Mournful’ focuses on making other people think you are sad. *Mournful* music makes listeners sad. So *mournful* does not talk much about someone’s own sadness, but more about making other people think you are sad.

Now we come to the top 50 collocates of *mournfully* (236 hits). As expected, *mournfully* has a semantic preference for physical action verbs, like *say/add/recite* (23/23/5), *gaze* (8), *look/stare* (7/5), *sign* (6), *sit* (4) etc. So the data support the claim I made earlier that *mournful(ly)* is more extroverted than introverted. Like *mournful*, *mournfully* is only modified by the downtoner *rather*.

5.2.6.2 Aidao

The BPC yields 90 hits of *mourn**. The search findings are given below:

Table 5.6. 1 Chinese equivalents of *mourn** in the BPC

<i>sang</i>	31 [<i>sangfu</i> 17 (including <i>mourning brooch</i> , <i>mourning bonnet</i> , <i>mourning clothes</i> , <i>mourning dress</i>), <i>fusang</i> 10, <i>baosang</i> 1, <i>shouxiao</i> 2, <i>huaquan</i> 1]
<i>zang</i>	5 [<i>songzangche</i>]
<i>beiai</i>	7
<i>aidao</i>	6
<i>shangxin</i>	6
<i>beishang</i>	5
<i>beitong</i>	4
Others	<i>qiliang</i> 2, <i>qiku</i> 1, <i>candan</i> 1 etc.

In 31 of 90 occurrences of *mourn**, the translation is related to the Chinese concept of *sang* (a broad concept related to dead people, including mourning, funeral etc), such as *sangfu* (mourning clothes), *fusang* (go into mourning), *baosang* (announce the death), *shouxiao* (in mourning), *huaquan* (mourning wreath). 5 of them are translated as *songzangche* (mourning coach), where *zang* is a Chinese concept meaning the burial ceremony. Most of the above-discussed instances contain *mourning*. For the remaining occurrences, 6 of them are translated as *aidao*, 7 as *beiai*, 6 as *shangxin*, 5 as *beishang* and 4 as *beitong*. Other translations include *qiliang*, *qiku*, *candan*, which are used to describe things or situations causing sad feelings. *Shangxin*, *beishang* and *beitong* are used to translate its sense ‘to feel sad especially about someone’s death’, as revealed in the following examples:

- 1) 现在他的躯体去和他的心会合了，我们就没有理由为他感到悲痛了。如果还感到悲痛(*beitong*)就太自私了。我爱他就像爱自己的父亲，所以才这样说。 Now that his body's gone to join his heart, I don't see that we got reason to *mourn*, unless we're pretty damned selfish, and I'm sayin' it who loved him like he was my own pa.
- 2) 思嘉也曾想把皮蒂姑妈找来，缓和一下她良心上的不安，但是她又犹豫了，姑妈要是来了也许会更糟，因为她对弗兰克的死由衷

地感到悲痛(*beitong*)。For a moment Scarlett thought of summoning Pittypat to stand between her and her conscience but she hesitated. Pitty would probably make matters worse, for she honestly *mourned* Frank.

The sense of showing sadness about someone's death is reflected in the Chinese translation equivalents related to funeral or other relevant communal or social practices, such as *sang-words*, *aidao* and *beiai*.

Beiai, *shangxin*, *beishang* and *beitong* have been discussed in earlier sections, so this section will only focus on *aidao*. Below is the search result for *aidao* (1402 hits) from the CCL corpus:

Table 5.6. 2 Significant collocates of *aidao* in the CCL corpus

Showing	Modifier	Cause	Object	Subject	Others
<i>baoshi</i> 832	<i>shenqie</i> 205	<i>yu'nan</i> 195	<i>sizhe</i> 46	<i>quanguo</i> 56	<i>jiashu</i> 126
<i>zhuanda</i> 11	<i>chentong</i> 124	<i>shishi</i> 189	<i>sinanzhe</i> 63	<i>renmin</i> 46	<i>qingshu</i> 63
<i>biaoda</i> 10	<i>chengzhi</i> 17	<i>buxing</i> 59	<i>li'nanzhe</i> 15	<i>tongbao</i> 25	<i>weiwen</i> 115
<i>xiabanqi</i> 8	<i>shen(biao)</i> 16	<i>qushi</i> 35	<i>shouhaizhe</i> 13	<i>zhengfu</i> 38	<i>beitong</i> 19
<i>jiangbanqi</i> 8		<i>yuhai</i> 17	<i>yuhai</i> 12	<i>zongtong</i> 37	<i>daonian</i> 9
<i>yandian</i> 31		<i>bingshi</i> 17		<i>zhuxi</i> 24	<i>huainian</i> 8
<i>fangshi</i> 45		<i>shenwang</i> 12		<i>lingdaoren</i> 17	
<i>xingshi</i> 14		<i>xisheng</i> 10		<i>zongli</i> 10	

From the table, it can be seen that *aidao* is 'to show one's regret or sadness about someone's death', which is clearly reflected in its verbal collocates in the first column: *biaoshi/biaoda* (express), *zhuanda* (convey). It is a social practice to convey the whole nation's sadness about someone's death, via *jiang banqi/xia banqi* (flying the flag at half mast). National leaders usually communicate their sadness about a very important person's death or a big group of people killed in other countries via telegram (*yandian*, *zhidian*). So it is mainly a matter of *xingshi* (form) and done in certain *fangshi* (ways). The second column lists its premodifiers, such as *shenqie/shen (biao)* (deep/deeply), *chentong* (with grief) and *chengzhi* (sincere). The third column gives the causes for

aidao, such as *yu'nan/yuhai/shenwang* (be killed), *shishi/qushi* (die), *buxing* (misfortune), *bingshi* (die of a disease). The fourth column is concerned with the objects of *aidao*, like *sinanzhe/li'nanzhe/yuhaizhe/shouhaizhe* (victim). The subjects of *aidao* include *quanguo* (the whole nation), *renmin* (people), *tongbao* (compatriot), *zhengfu* (government), *zongtong* (president), *zhuxi* (chairman), *zongli* (prime minister), *lingdaoren* (leader) etc. The final column gives its other collocates, such as *qinshu/jiashu* (family members), *weiwen* (condolence), *beitong* (grief) etc. *Aidao* is often done together with *weiwen*. The corpus data listed here is more related to a public mourning, associated more with a nation for the victims of a disaster. All of the objects of *aidao* listed here are killed because the CCL corpus contains a lot of newspaper articles, which often report social tragedies. Actually, *aidao* is also used to talk about privately mourning a family member, as exemplified in the following example from the CCL corpus:

- 3) 人死后亲属和亲戚中的晚辈或平辈穿孝服, 表示[哀悼](*aidao*)
 ‘After someone dies, his/her family members and relatives of the same or junior generation will wear *mourning* clothes to show *aidao*’.

5.2.6.3 Comparison

Different from *mourn*, *aidao* is a word used only to show one's sorrow about someone's death, based on the social or communal practice. *Aidao zhe* (*mourner*) is often *shangxin* or *beitong*, but not necessarily so. *Mourn* can be used to talk about one's feeling of sadness and showing of sadness. In addition, it can be used in a metaphorical way to talk about the feeling of loss of something, which is absent in *aidao*'s meaning. *Aidao* is predominantly used to talk about the showing of someone's sorrow or regret about the

death of someone. Chinese words distinguish clearly between the feeling and showing of sadness. English words, however, do not distinguish them sharply. Contexts have to be depended on to decide whether you are talking about communal or culturally-mediated behaviour or someone's internal feeling.

5.2.7 *Depression** & *Melancholy** and their Chinese equivalents

5.2.7.1 *Depression** & *Melancholy**

5.2.7.1.1 *Depression**

This section will look at *depress*, *depression*, *depressing* and *depressingly*, leaving the discussion of *depressed* to the next section, where it can be compared and contrasted with *melancholy* (adj.) For the discussion of *depress*, let us start with its dictionary definitions:

LDCE: *v* [T]

▢ to make someone feel very unhappy ▢ to prevent an economy from being as active and successful as it usually is ▢ *formal* to press something down, especially a part of a machine ▢ *formal* to reduce the value of prices or wages

OALECD: *v* [Tn]

▢ (sb) sad and without enthusiasm 使 (某人) 忧愁, 消沉, 沮丧 ▢ press, push or pull (sth) down 将 (某物) 压下, 推下, 拉下 ▢ make (esp trade) less active 使 (尤指贸易) 不活跃, 不景气, 萧条:

COBUILD: [VERB] Vn

1 If someone or something **depresses** you, they make you feel sad and disappointed.

2 If something **depresses** prices, wages, or figures, it causes them to become less.

LDCE's definition 'to make someone feel very unhappy' offers nothing valuable.

OALECD says that *depress* implies 'without enthusiasm' and COBUILD uses 'disappointed' instead.

For its corpus data, its top 50 collocates based on 4745 hits (inclusive of *depressing* and *depressed*) show that *be* (1361) tops the collocate list, which is because, in most cases, *depress* is used as a past participle (some of them are actually adjectives, but the tagging tool of the BOE tags them as VBN). It has a strong colligation with prepositions, mainly *by* (252), *about* (129), *for* (100) and *at* (76). Its word sketch reveals that most of its subjects are associated with its feeling sense, such as *prediction* (23), *thought* (11), *feeling* (9), *war* (7), *event* (6), *death* (5), *news* (5), *fact* (4) *memory* (4), *alcohol* (4). Its modifier sketch tells us that *depress* are often related to *clinically* (52), *chronically* (11), *suicidally* (7) and *terminally* (4), which are all associated with its unhealthy condition. Its modifiers show that ‘*depress*’ could be very serious since it can be intensified by *severely*(50), *profoundly*(19), *deeply* (57), *terribly*(8), *thoroughly*(10), *utterly* (8), *seriously* (22), *extremely*(18), *very* (143), *incredibly*(4), and less serious because it can also be modified by *somewhat* (7), *slightly*(12) and *mildly* (4), and that it could happen *frequently*(7), *temporarily* (6) or *permanently* (4).

Now let us look at its nominal form *depression*, whose dictionary entries are shown as follows:

LDCE:

▣[U and C] **a)** a medical condition that makes you very unhappy and anxious and often prevents you from living a normal life **b)** a feeling of sadness that makes you think there is no hope for the future ▣**the (Great) Depression:** the period during the 1930s when there was not much business activity and not many jobs ▣[U and C] a long period during which there is very little business activity and a lot of people do not have jobs →recession ▣[C] a part of a surface that is lower than the other parts ▣[C] *technical* a mass of air under low pressure, that usually causes rain

OALECD:

①[U] being *depressed*; low spirits 忧愁; 沮丧; 消沉 ②[C] hollow sunken place in the surface of sth, esp the ground; dip 凹陷处; (尤指) 洼地, 坑: ③[C] period when there is little economic activity, and usu poverty and unemployment 经济萧条期. ④[C] (a) (winds caused by a lowering of atmospheric pressure 气压降低 (形成的气流) . (b) area where this happens 低气压区.

COBUILD:

1 [N-VAR] **Depression** is a mental state in which you are sad and feel that you cannot enjoy anything, because your situation is so difficult and unpleasant.

2 [N-COUNT] A **depression** is a time when there is very little economic activity, which causes a lot of unemployment and poverty.

3 [N-COUNT] A **depression** in a surface is an area which is lower than the parts surrounding it.

4 [N-COUNT] A **depression** is a mass of air that has a low pressure and that often causes rain.

For the meaning associated with the mental state, LDCE breaks it down into two senses:

one refers to ‘a medical condition that makes you very unhappy and anxious and often prevents you from living a normal life’ and the other is ‘a feeling of sadness that makes you think there is no hope for the future’, which is quite similar to OALECD’s ‘without enthusiasm’ and ‘low spirits’, or COBUILD’s ‘disappointed’ and ‘you cannot enjoy anything’. Now let us turn to its collocates based on 14247 instances in the BOE. Its top 50 collocates reveal that it has strong colligations with the definite article *the* (1602), prepositions, predominantly *of* (1910), *in* (660), *for* (503), *from* (454), *into* (156), *by* (122), *about* (86), and *be* (857). Based on its word sketch, the verbs of *depression* include *suffer* (185), *treat* (146), *cause* (110), *alleviate* (32), *associate* (23), *diagnose* (20), *lift* (19), *trigger* (17), *underlie* (17), *worsen* (16), *deepen* (15), *combat* (15), *cure* (12), suggesting semantic preferences for causing, suffering, worsening, lessening or curing *depression*. Although some of them can be used in a metaphorical sense, most of them are associated with a medical condition. This is supported by its modifiers: *post-*

natal/postnatal/Post-natal (148/111/15), postpartum (126), manic (161), clinical (180), ST (38), tropical (61), chronic (72), maternal (26), suicidal (22). Its degree modifiers include *severe (288), deep (208), moderate (29), shallow (23).* *Deep* and *shallow* show that *depression* is understood in terms of the metaphorical image: A FLUID IN A CONTAINER, with its depth correlated with its intensity. The following table shows the words frequently used in the ‘and/or’ pattern for *depression* (the figures following the word show the frequencies and the figures after the frequency indicate the salience):

Table 5.7. 1 And/or pattern for *depression* in the BOE

<u>and/or</u>	<u>Fr</u>	Salience	<u>and/or</u>	<u>Fr</u>	Salience	<u>and/or</u>	<u>Fr</u>	Salience
anxiety	<u>596</u>	9.78	stress	<u>134</u>	7.24	headache	<u>47</u>	6.65
irritability	<u>52</u>	8.18	loneliness	<u>36</u>	7.22	anger	<u>82</u>	6.62
schizophrenia	<u>75</u>	8.17	lethargy	<u>26</u>	7.22	sadness	<u>27</u>	6.42
alcoholism	<u>40</u>	7.51	self-esteem	<u>39</u>	7.06	nausea	<u>20</u>	6.4
fatigue	<u>64</u>	7.48	despair	<u>47</u>	7.01	exhaustion	<u>21</u>	6.35
mania	<u>32</u>	7.36	anorexia	<u>20</u>	6.79	addiction	<u>30</u>	6.33
insomnia	<u>31</u>	7.35	paranoia	<u>23</u>	6.76	diabetes	<u>32</u>	6.32
psychosis	<u>30</u>	7.34	asthma	<u>31</u>	6.7	dementia	<u>17</u>	6.29
disorder	<u>153</u>	7.28						

From the above table, it can be seen that the list includes a lot of diseases or disease-related terms, such as *schizophrenia, alcoholism, mania, insomnia, psychosis, disorder, anorexia, paranoia, asthma, headache, nausea, addiction, diabetes, dementia* etc. It further supports the earlier claim that *depression* as a noun is more used as ‘a medical condition’. Other words in the list are mostly related to feelings, such as *anxiety, irritability, stress, loneliness, lethargy, despair, anger, sadness.* One noteworthy word is *self-esteem*, which means ‘good opinion of one’s own character and abilities’ (OALECD). It occurs 39 times in this pattern, probably suggesting that *depression* is a condition that makes one lose one’s self-esteem and have a low opinion of one’s own character and abilities. We find *fatigue* and *exhaustion* in the list. *Fatigue* is ‘a feeling of

extreme physical or mental tiredness' (OALECD). So this might allude to the fact that *depression* arises from mental tiredness. *Anxiety* tops the list, so we have the reason to argue that the mental tiredness might result from anxiety. Of course, the feeling of *depression* might also incorporate other feelings, such as *anger*, *sadness*, *despair*, *loneliness*. Consider the following paraphrases of depression extracted from the BOE:

- 1) Remember that **depression is** merely anger turned in against yourself, instead of being directed outwards.
- 2) ...for others the **depression is** a symptom of problems with life rather than an illness, and thus not curable by drug treatment. **Depression is** a painful experience...
- 3) Psychologists have sometimes argued that **depression is** more concerned with our feelings about the past, while anxiety is to do with fears for the future, which when you think about it, seems to make a lot of sense. DEPRESSION? Life events (for a list of these see the Holmes-Rahe scale in the section on stress). Psychological factors. Some psychologists attribute depression to 'learned helplessness', when we feel so out of control of our lives and it seems pointless to even try. Others attribute depression to 'faulty thinking', i. e. the way we view the world affects our moods. Psychoanalytical theories see depression as being a result of aggression being turned inwards upon ourselves.
- 4) 'For me, the **depression is** connected to the gap between our expectations and our views of what matters in life, and the frustration which comes with never really achieving those.'

Depression is a well-researched topic in psychology, which explains why we can find many paraphrases in the corpus. According to the above examples, *depression* seems to be anger turned against oneself and this supports the earlier claim that it has to do with poor self-esteem; it is correlated with disappointment or despair, depending on its seriousness. Example 3 tells us that members of a discourse may disagree about what a discourse object might mean. For some psychologists, *depression* is more concerned with the feelings about the past, but the corpus data

appear to show that it is not quite the case because anxiety, which is to do with fears about the future, tops the list in Table 5.7.1.

2418 hits of *depressing* as an adjective are found in the corpus. Its word sketch shows that *depressing* can be modified by amplifiers, such as *very* (91), *deeply* (25), *really* (22), *profoundly* (9), *thoroughly* (8), *terribly* (7), *utterly* (5) *extremely* (5), *downright* (5), *totally* (5), *immensely* (5), and downtoners, like *pretty* (25), *rather* (27), *little* (8), *merely* (4), *quite* (11), etc. For the words it modifies, almost all of them are related to the sense of feeling: *account* (62), *reading* (40), *effect* (39), *news* (30), *picture* (30), *thought* (28), *experience* (26), *finale* (26), *run* (22), *statistic* (14), *aspect* (12), *defeat* (16), *sight* (14), *reality* (9), *spectacle* (9), *reminder* (8), *conclusion* (7), *truth* (7), *tale* (7), *prospect* (6), *regularity* (6), *scenario* (5), *episode* (5), *lack* (5), *storylines* (5). It seems that many of these words can be modified by *sad* and *depressing* is, in many cases, only a stronger version of *sad* in the sense of ‘making one feel unhappy’.

For *depressingly* (310 hits), its top 50 collocates show that it semantically prefers neutral or even positive adjectives which have no obvious negative inclinations, like *familiar* (56), *predictable* (12), *similar* (9), *clear* (5), *large* (5), *inevitable* (4), *long* (4). It seems that *depressingly* has lost its ‘making sad’ flavour and developed into a complete degree adverb, without negative tendency. This is referred to as ‘delexicalisation’ (Sinclair, 1996). Similar cases would be *awfully*, *terribly*, *woefully* etc.

5.2.7.1.2 *Melancholy**

This sub-section will present a comparative analysis of *melancholy* and *depressed*, followed by the discussion of *melancholic*. To begin with, let us look at the dictionary entries of *depressed* and *melancholy*:

LDCE:

depressed *adj*

▣**a**) very unhappy **b**) suffering from a medical condition in which you are so unhappy that you cannot live a normal life ▣**▶**an area, industry etc that is depressed does not have enough economic or business activity ▣**▶** *formal* a depressed level or amount is lower than normal *a depressed appetite*

melancholy

adj very sad

n [U] [Origin: *melancolie*, from *Late Latin* *melancholia*, from *Greek*, from *melas* 'black' + *chole* 'bile'] *formal* a feeling of sadness for no particular reason

OALECD:

depressed *adj*

sad and without enthusiasm 忧愁的; 消沉的; 沮丧的:

melancholy

n [U] (tendency towards) deep sadness which lasts for some time;

depression 忧郁; 抑郁.

adj (a) very sad; depressed 悲哀的; 沮丧的 (b) causing sadness 使人忧郁的

COBUILD:

depressed

1 [ADJ] usu v-link ADJ If you are **depressed**, you are sad and feel that you cannot enjoy anything, because your situation is so difficult and unpleasant.

2 [ADJ] A **depressed** place or industry does not have enough business or employment to be successful.

melancholy

1 [ADJ] You describe something that you see or hear as **melancholy** when it gives you an intense feeling of sadness.

2 [N-UNCOUNT] **Melancholy** is an intense feeling of sadness which lasts for a long time and which strongly affects your behaviour and attitudes. (LITERARY)

3 [ADJ] If someone feels or looks **melancholy**, they feel or look very sad. (LITERARY)

For *depressed* as a feeling, LDCE's explanation of 'very unhappy' is meaningless. OALECD's 'without enthusiasm' and COBUILD's 'you cannot enjoy anything' can show that getting *depressed* means one has lost the keen interest in life. For *melancholy*, LDCE claims that it is 'a feeling of sadness for no particular reason'. OALECD says the salient feature for *melancholy* is 'deep sadness which lasts for some time' and it uses *depression* to define *melancholy*. COBUILD defines *melancholy* as 'an intense feeling of sadness which lasts for a long time and which strongly affects your behaviour and attitude'. So what is shared by all three dictionaries is 'a feeling of deep/intense sadness'. 'A feeling which lasts for some time' is only shared by OALECD and COBUILD.

To determine their other distinguishing features, we need to look at the corpus data. As mentioned earlier, many instances of *depressed* which are actually adjectives are tagged as VBN. However, we cannot differentiate them from the real past-participle cases using a corpus technology and merge them into this group, so we can only analyse what the BOE software program provides for us, albeit its incompleteness. In the BOE, *melancholy* only occurs 16 times as a noun, including *melancholies*. There are too few instances to make a claim, so I will focus my discussion on *melancholy* as an adjective.

The BOE contains 3700 hits of *depressed* (*adj.*) and 1772 hits of *melancholy* (*adj.*). For the sake of space, I only summarise the relevant findings as follows:

Depressed (3700):

Colligations:

a) link verbs: be (171), feel (183), become (191), get (151), remain (49), look (30), seem (31)

b) preposition: about (55)

Semantic preferences:

a) feeling-related: *patient* (76), *mood* (43), *person* (78), *mother* (54),
state (54)

b) non-feeling: *market* (97), *area* (87), *price* (75), *economy* (29)

Melancholy (1772):

Colligations:

a) link verbs: *feel* (13), *sound* (10)

b) preposition: *about* (25)

Semantic preferences:

a) thinking: *thought* (11), *musings* (3), *reflection* (3)

b) music: *music* (13), *song* (11), *note* (7), *tone* (5), *vocal* (3),
melody (3),

c) things making one sad: *ending* (6), *tale* (5), *occasion* (5), *sight* (4),
truth (8), *task* (6), *reading* (4)

d) things related to feeling or showing sadness: *mood* (14), *soul* (5),
smile(5), *expression* (4)

The above figures reveal that *depressed* has a much stronger colligation with link verbs, especially *feel* (49% > 7%, taking their total numbers of occurrences in the corpus into consideration) than *melancholy*. In terms of semantic preferences, *depressed* tends to modify a person or an emotional state, while *melancholy* is more likely to modify words related to thinking or music. The above figures seem to suggest that ‘*depressed*’ focuses more on the inner emotional state (*feel*, 183), but it is also often outwardly manifested (*look*, 30; *seem*, 31) and is likely to be perceived. The use of ‘sound melancholy’ and the semantic association of music as mentioned above seem to suggest that a *melancholy* feeling tends to be outwardly manifested, especially via voice or music. Its semantic association with thinking shows that it is a typical inner state as well. In addition, *melancholy* also modifies positive words, like *comedy* (6) and *pleasure* (4).

Significant collocates for these two words in other patterns include:

Unique patterns:

Depressed:

Modifier: *economically* (25), *clinically* (21), *severely* (34), *deeply* (21),
seriously (11), *extremely* (11), *terribly* (6), *mildly* (6)

And/or: *anxious* (110), *angry* (44), *lonely* (38), *tired* (23), *miserable* (21), *irritable* (17), *lethargic* (12), *frustrated* (10), *frightened* (10)

Melancholy:

Modifier: *deep* (8)

And/or: *romantic* (10), *sweet* (8), *profound* (8), *beautiful* (7), *morbid* (5), *gentle* (5), *exquisite* (4)

Common patterns (*depressed/melancholy*):

Modifier: *little* (7/10); *rather* (12/11)

Obviously, *economically* and *clinically* come from the non-feeling senses of *depressed*.

It is clear from the above list that *melancholy* is only modified by *deep*, (*a*) *little* and *rather*, unlike *depressed* which is intensified by *severely*, *deeply*, *seriously*, *extremely*, *terribly*, *mildly*, *rather* and (*a*) *little*. It is interesting that all the words used in the ‘and/or’ pattern for *depressed* are negative, especially the highly frequent ones, i.e. *anxious*, *angry*, *lonely* and *tired*, which constitute the meaning of *depressed*. In contrast, almost all the words in the list for *melancholy* are positive, except *morbid*. The following concordance listing¹⁹ can show that all of the instances of *melancholy* come from the ‘and’ rather than the ‘or’ pattern, indicating that *melancholy* has a positive flavour in meaning and is thus less negative in terms of its semantic prosody:

and early nineteenth-century wheals have a wilderness, full of briars and romantic	melancholy	, romantic feel in ruin that they doubtless
pensive and thoughtful shepherdess, alone in a	melancholy	, and by visiting it shortly before a burial
prankster and callous joker, and Pierrot, the	melancholy	romantic landscape that is painted dreamily
to be this star. People think I'm really	melancholy	romantic, and we find even more options
Fenner's gouaches a grandly romantic, faintly	melancholy	and romantic and all whispery. I'm not
south-facing slopes. <p/> It is a slightly	melancholy	quality somewhere in the same emotional
There is even, perhaps, a touch of romantic	melancholy	, romantic place, where you half expect
Ingres suggested Franet's romantic, somewhat	melancholy	about the whole thing. And elsewhere the
is like every Wong Kar-wai character: a	melancholy	temperament by silhouetting his head against
		romantic. And he has the bruises to prove

My inspection of its other positive collocates in their concordances confirms this claim, though, in some instances, it is used as a noun rather than an adjective. Due to the

¹⁹ In Word Sketch, by clicking on the frequency after a collocate, you will get all the relevant instances.

inaccurate tagging, they are misplaced here. However, it does not affect our observation that *melancholy* is less negative. Look at the following example:

- 5) His mind was not political, only contemplative. That contemplation is pessimistic; perhaps **melancholy is** a better word.
- 6) In Iran and Sri Lanka for example a degree of **melancholy is** taken as an indication of a person's depth.

The above examples suggest that *melancholy* is closely associated with contemplation, which can be taken as an indication of a person's depth. The correlation of contemplation with depth should also apply in the English culture, not only in Iran and Sri Lanka, as Example (6) shows.

For *melancholic*, the dictionaries offer the following:

LDCE:

adj literary feeling very sad

OALECD:

adj (having a tendency to be) melancholy 忧郁的; 抑郁的

COBUILD:

[ADJ] If you describe someone or something as **melancholic**, you mean that they are very sad. (LITERARY)

Both LDCE and COBUILD say that *melancholic* is literary, which seems to be the very thing that differentiates it from *melancholy*. 209 hits of *melancholic* as an adjective are found in the corpus. Due to its low frequency in the corpus, we can only find a few meaningful words it modifies in its top 50 collocates, like *tone* (5), *song* (4), *mood* (3) and *music* (3). It appears, from its collocates, that *melancholy* and *melancholic* are very much similar, only differing in that the latter is more literary. *Sanguine* (8) is particularly prominent in the collocates of the 'and/or' pattern, suggesting the antonym of *melancholic* seems to be 'hopeful, optimistic'. *Melancholic temperament* (3) is also found in its word sketch, showing that *melancholic* has a long-lasting quality which

may constitute one's character. *Gorgeously, majestically, beautifully, sweetly, wonderfully* in its modifiers seem to show that *melancholic*, albeit it means sad or pessimistic, is not a completely negative word since it can be modified by words with positive connotations.

5.2.7.2 Chinese equivalents of *depression** & *melancholy**

60 hits of *depress** and 54 hits of *melancholy** are found in the BPC. Their most frequently occurring Chinese translation equivalents are given below:

Table 5.7. 2 Top 3 Chinese equivalents of *depression** & *melancholy** in the BPC

	<i>Depression*</i>	<i>Melancholy*</i>
Chinese translational equivalents	<i>jusang</i> : 15 <i>yayi</i> : 5 <i>qinhxu diluo</i> : 4	<i>youyu</i> : 21 <i>shanggan</i> : 3 <i>beishang</i> : 3

We have discussed *beishang* in an earlier section and will focus on *jusang, yayi, qingxu diluo, youyu* and *shanggan* in this section.

5.2.7.2.1 *Jusang*

Jusang occurs 1904 times in the CCL corpus and its top collocates are given below:

Table 5.7. 3 Significant collocates of *jusang* in the CCL corpus

Introversion/ extroversion	Modifier	Feeling	Others
<i>gandao/gan</i> 164/25	<i>hen</i> 106	<i>shiwang</i> 41	<i>ling</i> 135
<i>juede/ganjue</i> 41/12	<i>shifen</i> 66	<i>tongku</i> 24	<i>lingren</i> 63
<i>xinli/xinzhong</i> 21/17	<i>feichang</i> 55	<i>wanfen</i> <i>fennu</i> 21	<i>shi</i> 118
<i>jingshen</i> 45	23	<i>juewang</i> 18	<i>shuo</i> 168
<i>qingxu</i> 85	<i>ji/jile</i> 16/17	<i>xiaochen</i> 13	<i>xiang</i> 58
<i>xinqing</i> 77	<i>jidu/jiwei</i> 11/11	<i>youyu</i> 13	<i>shibai</i> 30
<i>xiande</i> 25	<i>yidian/youdian</i> 13/11	<i>beiai</i> 13	<i>xiaoxi</i> 22
<i>shenqing</i> 95	<i>youxie/yixie</i> 22/11	<i>gudu</i> 12	<i>yizhen</i> 10
<i>shense</i> 16		<i>shangxin</i> 11	
<i>lianse</i> 14		<i>jiaolv</i> 10	
<i>biaoqing</i> 13			
<i>lian/lianshang</i> 27/10			
<i>manlian</i> 11			
<i>shengyin</i> 10			

What we can see from the collocates in the first column is that *jusang* is an emotion both experienced as an inner state, as shown in *gandao/gan/juede/ganjue* (feel), *xinli/xinzhong* (in the heart), *jingshen* (spirit), *qingxu* (mood), *xinqing* (frame of mind), and involving outward manifestations, as indicated in *biaoqing/shenqing/shense/lianse* (facial expression), *lian/manlian/lianshang* (face/full face/on the face), *shengyin* (voice). It can be modified by *hen/shifen/feichang* (very), *wanfen/jile/ji/jidu/jiwei* (extremely), *yidian/youdian/yixie/youxie* (a little). It often co-occurs with *shiwang* (disappointment), *tongku* (anguish), *fennu* (indignation), *juewang* (despair), *xiaochen* (low-spirited), *youyu* (melancholy), *beiai* (sorrow), *gudu* (loneliness), *shangxin* (sadness), *jiaolǚ* (anxiety) and *bu'an* (uneasiness). It also tends to co-occur with *ling/lingren/shi* (cause/make), *shuo* (say), *xiang* (think), *shibai* (failure), *xiaoxi* (news) and *yizhen* (lasting for a short time). *Shuo* and *xiang* seem to support the claim I made earlier that *jusang* is both extroverted and introverted. *Shibai* is the main cause for *jusang*. *Yizhen* (lasting for a short time) seems to suggest that *jusang* might come suddenly and go quickly. It is an immediate reaction to what has just happened, such as failure. It is specific because it has a definite reason. It may affect somebody for a long time, depending on how serious it is, but more often, it is a strong emotion which only lasts for a short time and fades away soon.

5.2.7.2.2 *Yayi*

Yayi is formed by two synonymous verbs: *ya* and *yi*, both meaning ‘press down’. In XHCBL, it is glossed as:

对感情、力量等加以限制，使不能充分流露或发挥 (of feelings, strength, etc.)
constrain; inhibit; depress; hold back

This is only its definition as a verb and its sense as a feeling is not included in it. To find its meaning as a feeling, we have to resort to the CCL corpus. 2581 hits of *yayi* are found there and its top collocates are given below:

Table 5.7. 4 Significant collocates of *yayi* in the CCL corpus

Introversion	Modifier	Feeling	Others
<i>gandao/gan</i> 140/84	<i>yanzhong</i> 19	<i>chenzhong</i> 44	<i>shehui</i> 51
<i>ganjue/juede</i> 35/31	<i>jidu</i> 11	<i>tongku</i> 39	<i>shenghuo</i> 43
<i>ganqing/qinggan</i> 71/17	<i>feichang</i> 20	<i>fennu</i> 31	<i>qifen</i> 40
<i>xinqing</i> 71	<i>dai</i> 11	<i>kumen</i> 22	<i>renxing</i> 40
<i>qingxu</i> 63	<i>yidian</i> 8	<i>chenmen</i> 18	<i>ziwo</i> 39
<i>neixin</i> 68		<i>kongju</i> 18	<i>gexing</i> 36
<i>xinli/xinzhong</i> 47/42		<i>jiaolü</i> 14	<i>xingge</i> 11
<i>xintou</i> 25			<i>yizhi</i> 33
<i>sixiang</i> 21			<i>changqi</i> 47
<i>xinling</i> 13			<i>niuqu</i> 20
			<i>shufu</i> 19
			<i>baofa</i> 16

It can be seen that *yayi* is a word predominately indicating an inner emotional state, as shown in its collocates in the first column: *gandao/gan/ganjue/juede* (feel), *ganqing/qinggan* (feeling/emotion), *xinqing* (frame of mind), *qingxu* (mood), *neixin/xinli/xinzhong/xintou* (in the heart), *sixiang* (thought) and *xinling* (mind). The second column shows that it is often modified by the degree adverb *yanzhong* (severely), amplifiers, typically *feichang* (very), *jidu* (extremely), and downtoners, predominantly *yidian/dai* (slightly). It co-occurs with other negative feelings like *chengzhong* (heaviness), *tongku* (anguish), *kumen* (gloominess), *chenmen* (dreariness), *kongju* (fear) and *jiaolü* (anxiety). The last column reveals that *yayi* is closely related to one's life (*shenghuo*) and society (*shehui*), and that it is semantically associated with *qifen* (atmosphere), *renxing* (human nature), *ziwo* (self), *gexing* (personality), and *xingge* (character). From its semantic association with *gexing* and *xingge*, we can argue that it

has become a temperament which affects a person's way of thinking, feeling and behaving or part of the human nature. *Changqi* (for a long time) and *yizhi* (always/all the time) show that it is a state extended in time. In the modern society, people are usually chained (*shufu*) by duties, customs or traditions. Their desires and impulses are suppressed, but finally they might explode. Therefore, *yayi* is an unpleasant feeling when one is not satisfied with the present situation and wants to change it, but s/he cannot do so due to some constraints from the society or somewhere else.

5.2.7.2.3 *Qingxu diluo*

Qingxu diluo literally means 'qingxu (mood) is low'. Chinese conceptualises that if one feels low-spirited, his/her *qingxu* will get lower. It occurs 252 times in the CCL corpus. The corpus data shows that it often collocates with *jingshen*/mental (13), *yin* /because (13), *kunnan*/difficulty (7), *gandao*/feel (5), *yangzi*/appearance (5), *jusang*/dejected (5) and *yi* (depressed). Its collocates tell us that *qingxu diluo* is usually caused by difficulty. It is a word describing one's mood, so it is more experienced as an inner emotional state, as shown in *jingshen* and *gandao*. However, it also collocates with *yangzi*, which seems to suggest that one tends to have some outwardly manifested behaviour or facial expressions when one is *qingxu diluo*. It is a mild word to describe a person's sadness, disappointment and frustration. *Kunnan* is the usual and typical cause for the lowering of one's *qingxu*. Sometimes, failure or a setback will also lead to *qingxu diluo*, but usually the result will be more serious than that.

5.2.7.2.4 *Yiyu*

As mentioned earlier, the parallel corpus used for this research is not a contemporary one, which cannot reflect the recent linguistic change. Hence, the researcher's intuition has to help find the closest equivalent in the other language. For *depression*, there is a modern translation equivalent-- *yiyu* or *yiyuzheng* (medical condition). *Yiyu* is formed by two words: *yi* (press down) and *yu*. The XHCBL's glosses for *yu* and *yiyu* are given as follows:

- yu*: (忧愁、气愤等) 在心里积聚不得发泄 [(of sorrow, anger etc.) accumulate in the heart, but cannot release] (of sorrow, anger, etc.) pent-up; gloomy; depressed
yiyu: 心有愤恨, 不能诉说而烦闷 [People will feel annoyed if they cannot reveal their feelings to others when they bear resentment towards something] depressed; despondent; gloomy

The real meaning of *yiyu* is far more than is given in this dictionary. It is a quite modern concept which has been coined to refer to a very serious unhealthy state that affects a person's normal life, including behaviour, thinking or feeling. It can be said to be a very strong version of *depression* in the sense of 'feeling sad or unhappy'. In practice, *yiyu* is also used to talk about one's intense feeling of sadness. To capture its semantic subtleties, let us look at its collocates in the CCL corpus. 778 hits of *yiyu* are found in the corpus and its top collocates are shown below:

Table 5.7. 5 Significant collocates of *yiyu* in the CCL corpus

Introversion/extroversion	Feeling	Others
<i>jingshen</i> 68	<i>jiaolǚ</i> 33	<i>huanzhe/huan</i> 23/23
<i>qingxu</i> 57	<i>zao/fanzao</i> 13/13	<i>zhiliao</i> 17
<i>xinqing</i> 52	<i>bukuai</i> 13	<i>zhengzhuang</i> 16
<i>gandao</i> 14	<i>gudu</i> 12	<i>bingren</i> 10
<i>xinzhong/xinli</i> 10/9	<i>jinzhang</i> 10	<i>jibing</i> 10
<i>shenqing</i> 13	<i>tongku</i> 9	<i>shenjing</i> 10
	<i>xiaochen</i> 8	<i>shimian</i> 9
	<i>beiguan</i> 7	<i>si</i> 17
	<i>kumen</i> 7	<i>zisha</i> 12
	<i>beishang</i> 7	

What is revealed from the first column is that *yiyu* collocates predominantly with words of introversion, such as *jingshen* (mental), *qingxu* (mood), *xinqing*(frame of mind), *gandao* (feel) and *xinzhong/xinli* (in the heart). However, it also has physical manifestations, like *shenqing* (facial expression). The negative feelings it co-occurs with include *jiaolǚ* (anxious), *zao/fanzao* (agitated), *bukuai* (unhappy), *gudu* (lonely), *jinzhang* (nervous), *tongku* (anguished), *xiaochen* (low-spirited), *beiguan* (pessimistic), *kumen* (gloomy) and *beishang* (sorrowful). Among them, *jiaolǚ* and *zao/fanzao* stand out, suggesting that *yiyu* is mainly caused by anxiety and agitation. The last column lists its other collocates which may contribute to unravelling its meaning. *Huanzhe/bingren* (patient), *huan* (contract/suffer from), *zhiliao* (treatment), *zhengzhuang* (symptom) and *jibing* (disease) are all related to its sense as a medical condition. *Shenjing* (nerve) might suggest that *yiyuzheng* is to do with a nerve problem. *Shimian* (insomnia) is one of its main symptoms. *Yiyu* may lead to *si* (death), usually by *zisha* (commit suicide). The following paraphrase from the CCL seems to be a very good summary of this emotion:

- 7) 抑郁是一种过度忧愁和伤感的情绪体验，一般表现为情绪低落，心境悲观... ‘*Yiyu* is an emotional experience of excessive worry and sadness, characterised by a low mood and pessimism’...

5.2.7.2.5 *Youyu*

Youyu is a combination of *you* (worry) and *yu*, which was just discussed in the last section. XHCBL glosses it as follows:

忧伤,愁闷[worry+sadness, worry+closed] melancholy; heavyhearted;
dejected

Youyu occurs 2288 times in the CCL corpus and its top collocates are given below:

Table 5.7. 6 Significant collocates of *youyu* in the CCL corpus

Introversion/extroversion	Modifier	Feeling	Others
<i>xinqing</i> 78	<i>hen</i> 83	<i>gudu/guji</i> 28/8	<i>zheng</i> 150
<i>qingxu</i> 38	<i>feichang</i> 23	<i>jimo</i> 13	<i>bing</i> 29
<i>jingshen</i> 37	<i>shifen</i> 17	<i>tongku</i> 25	<i>huan</i> 44
<i>xinzhong</i> 15	<i>lǚedai</i> 19	<i>beishang</i> 21	<i>huanzhe</i> 16
<i>gandao/gan/jue</i> 45/15	<i>xie/youxie</i> 17/19	<i>beiai</i> 16	<i>ziji</i> 51
<i>gan</i> 21	<i>dian/yidian</i>	<i>shanggan</i> 20	<i>chenmo</i> 25
<i>lianshang/lian</i> 33/26	31/14	<i>fannaο</i> 13	<i>chensi</i> 24
<i>lianse/shense</i> 25/20	<i>dandan (de)</i> 18	<i>jusang</i> 13	<i>cangbai</i> 22
<i>shenqing</i> 74	<i>si</i> 17	<i>kongju</i> 12	<i>si</i> 20
<i>biaoqing</i> 25		<i>juewang</i> 12	<i>qizhi</i> 9
<i>yanjing</i> 77		<i>jiaolǚ</i> 12	<i>xingge</i> 16
<i>muguang/yanshen</i> 37/37		<i>chouchang</i> 11	<i>xiao/weixiao</i> 20/18
<i>xiande</i> 27		<i>xiaochen</i> 11	<i>meili</i> 14
<i>liulu</i> 19		<i>kumen</i> 14	<i>langman</i> 9
<i>luchu</i> 18		<i>chenmen</i> 13	<i>wangzi</i> 9
<i>shengyin</i> 21			<i>wenrou</i> 9
			<i>chongman</i> 30

The first column shows that *youyu* is both introverted, as revealed in its collocates related to introversion, characteristically *xinqing* (frame of mind), *qingxu* (mood), *jingshen* (mental), *xinzhong* (in the heart) and *gandao/gan/ganjue* (feel), and extroverted, as indicated in its following collocates: *lianshang/lian* (on the face/face), *lianse/shense/shenqing/biaoqing* (facial expression), *yanjing* (eyes), *muguang/yanshen* (expressions in one's eyes), *xiande* (appear), *liulu/luchu* (reveal), *shengyin* (voice). The second column lists its modifiers: *hen/feichang/shifen* (very), *luedai/youxie/xie/dian/yidian* (a bit/slightly), *dandan de* (faint), and *si* (a minute quantity). The negative feelings it often co-occurs with include *gudu/guji/jimo* (lonely), *tongku* (anguished), *beishang/beiai* (sorrowful), *fannaο* (annoyed and worried), *jusang* (dejected), *kongju* (fearful), *juewang* (desperate), *jiaolǚ* (anxious), *shanggan/chouchang* (sentimental), *xiaochen* (low-spirited), *kumen* (gloomy) and *chenmen* (dreary). For the last column, *zheng/bing* (disease), *huan* (contract) and *huanzhe* (patient) clearly show that *youyu* is also associated with an unhealthy condition. *Ziji* (33+18) co-occurs with

youyu 51 times, which seems to suggest that *youyu* is something one bears alone. *Chenmo* (silence) and *chensi* (contemplation) reveal that *youyu* is associated with quietness. *Cangbai* (pale) might mean that *youyu* is related to a pale face. *Youyu* may also lead to death (*si*) if it is a serious condition that needs medical treatment. It is more an emotional state closely related to the less active attitude toward life and the world, slightly fatalistic. A *youyu* person may be resigned to an unpleasant situation or fact and think s/he can not change the situation, and therefore there is no point in trying. *Youyu* might be the basic nature and is shown in the way in which one reacts to situations or other people, so it is often linked with *xingge* (character) or *qizhi* (temperament). Look at the following examples in which *youyu* as a temperament or character is mentioned:

- 8) 不知道你自己发现没有，你的气质里有一种[忧郁](*youyu*)的东西。我喜欢 [忧郁](*youyu*)，我这个人也常常[忧郁](*youyu*) ‘I don’t know whether you’ve found that there is a vein of *melancholy* in your **temperament**. I like *melancholy* and I often feel *melancholy*.’
- 9) 从幼小时就失去了家庭的温暖，一生穷愁潦倒，形成了他[忧郁](*youyu*)、孤独、愤世嫉俗和敏感的个性 ‘He has lost the warmth from family and been very poor since he was very young, which helped to form his **melancholy**, lonely, cynical and sensitive **character**’.

Youyu is often semantically associated with *xiao/weixiao* (smile), *meili* (beautiful), *langman* (romantic), *wenrou* (gentle) and *wangzi* (prince), suggesting that sometimes it is totally positive. Consider the following examples:

- 10) ‘好啊，’他还是[忧郁](*youyu*)地 笑笑
“‘Ok,” he gave a *melancholy* smile.’
- 11) 而最漂亮的还是她那双[忧郁](*youyu*)的嫣然动人的眼睛
‘What is the most beautiful is her *melancholy* and enchanting eyes.’

Like *youshang*, the reason why *youyu* is less negative is probably because it contains the character *you*, a concept closely associated with intellectuals who were so intelligent

and knowledgeable that they could foresee what would happen to the people and country in the future, usually something bad. They were not happy about enjoying the temporarily peaceful life, but were always worried about the future of the country. Therefore, *you* was a Chinese concept correlated with knowledge, wisdom, thinking, farsightedness and worry. Its positive flavour still remains in some modern Chinese words and *youyu* is one of them. A degree of *youyu* is seen as an indicator of a person's depth.

Chongman (be filled with) is a typical verb that is used to describe *youyu*, as well as many other emotions. This is similar to the description of English emotions. In English, if you are filled with an emotion, or if it fills you, you feel it very strongly. We can say someone is filled with *admiration/joy/happiness/horror/fear/anger/doubt/remorse* etc. Therefore, Chinese and English are quite similar in the way emotions are conceived of, but they differ in specific conceptualisations. For example, in Chinese, *chongman* is only related to some emotions, not all. As discussed earlier, in Chinese, we use *chenjin* (be immersed into) to describe *beitong* with great intensity and we never use *chongman* to describe *gaoxing* (delight) or *shengqi* (annoyance). Sometimes, we need to add some other words to express the idea, such as 'chongman jingpei zhi qing' (be filled with admiration), where *zhi* (particle, used between attribute and head) and *qing* (feeling) are added to make the expression more like Chinese. Therefore, it seems that even if the two languages are similar in the metaphorical mapping between the source domain and the target domain, they differ in specific expressions and details. We will return to the metaphorical issue in detail later.

5.2.7.2.6 *Shanggan*

For *shanggan*, XHCBL offers the following:

因感触而悲伤 [feel sad due to emotional stirrings] sentimental; sick at heart: 对景思人,无限伤感。 Feel extremely sorrowful to see the sight and think of one's people

Shanggan is a specific kind of sadness particularly when you see the same place and find that the people you are close to do not stay there any longer. Things will change with time, but people tend to mourn those which used to exist. Hence, *shanggan* is the sadness brought about by emotional stirrings. 1240 hits of *shanggan* are found in the CCL corpus and its top collocates are shown below:

Table 5.7. 7 Significant collocates of *shanggan* in the CCL corpus

Introversion/extrovSION	Modifier	Feeling	Others
<i>qingxu</i> 41	<i>hen</i> 63	<i>youyu</i> 21	<i>ziji</i> 34
<i>xinzhong</i> 18	<i>shifen</i> 23	<i>tongqing</i> 9	<i>chongman</i> 19
<i>gandao/juede</i> 16/14	<i>feichang</i> 15	<i>gudu</i> 7	<i>ling/shi/jiao</i> 39/39/12
<i>shengyin</i> 12	<i>xie/youxie</i> 38/25	<i>bei</i> 7	<i>yinwei</i> 30
<i>yanjing</i> 10	<i>yixie</i> 11		<i>ynqi</i> 16
<i>yangzi</i> 10	<i>dian</i> 23		<i>xiang</i> 41
<i>shenqing</i> 8	<i>dai/luedai</i> 37/13		<i>qilai</i> 24
<i>liulu</i> 10	<i>si</i> 10		<i>bian</i> 19
<i>kan</i> 31			<i>zuopin</i> 14
			<i>huaijiu</i> 12
			<i>gushi</i> 11
			<i>huiyi</i> 11

In the above table, we find *qingxu* (mood), *xinzhong* (in the heart) and *gandao/juede* (feel), showing that *shanggan* is an emotion experienced as an inner state; we also find *shengyin* (voice), *yanjing* (eyes), *shenqing* (facial expression), *yangzi* (appearance), *liulu* (reveal), *kan* (see), suggesting that *shanggan* is easily visible and perceivable. We also find that its modifiers include *hen/shifen/feichang* (very), *xie/youxie/yixie/dian/dai/luedai* (slightly/a little) and *si* (a minute quantity). It often co-occurs with *youyu* (melancholy), *tongqing* (sympathetic), *gudu* (lonely) and *bei* (sad). Of them, *youyu* is

the salient one which shares a lot in common with *shanggan* in the sense that both are associated with people who are easily affected by sadness. For the last column, *chongman* suggests that, like many other emotions, *shanggan* is also understood in terms of gas or liquid. *Ziji* seems to suggest that *shanggan* is a feeling often confined to oneself. Things that can arouse a feeling of *shanggan* include *zuopin* (work of art), *gushi* (story) and *huiyi* (memory). From *ling/shi/jiao* (cause/make), *yinwei* (because) and *yinqi* (give rise to), it can be seen that *shanggan* is definitely triggered by something. *Qilai* (used after a verb or adjective to indicate the beginning or continuation of an action) and *bian* (thus/so) show that *shanggan* is caused by something that has just happened. *Xiang* (think) shows that *shanggan* is often the outcome of thinking of something. It is associated with *huaijiu* (nostalgic) because one feels affectionately about the experiences one had in the past. Actually, *shanggan* is a feeling one will have when s/he thinks of something worth-recalling but gone. S/he wants to have it again or wants to change it back to its original condition, but s/he feels so helpless. Like *beishang*, *shanggan* is also a concept associated with the fatalistic view, characterised by a sense of powerlessness due to the laws of nature and society.

5.2.7.3 Comparison

So far, we have discussed, in great detail, *depressed** & *melancholy** and their Chinese equivalents. It seems that *depressed* has a wider range, covering the Chinese words *yiyu*, *jusang*, *yayi* and *qingxu diluo*. Of course, strictly speaking, *qingxu diluo* is not a word but a phrase into which modifying elements, like *feichang/shifen* (very), can be inserted. This phrase is similar to *low-spirited* in the sense that both are conceptualised in terms

of DOWN. *Depressed* can be preceded by modifiers indicating various degrees, showing that it has a broad semantic coverage. Among its Chinese translation equivalents, *qingxudiluo* is the least intense and *yiyu* the most intense. *Jusang* and *yayi* emphasise different aspects. The former focuses on disappointment, frustration and hopelessness and the latter on the difficulty or impossibility of untying the knots and the incompetence to release one's sad feelings. *Depressed* and *yiyu* correspond to each other in the sense that they both denote a serious emotional state of sadness and a medical condition. The same applies to *melancholy* and *youyu*. Both *youyuzheng* and *yiyuzheng* can refer to a medical condition and may be preferred by different groups of people. *Yiyuzheng* seems to be more formal, more academic and therefore occurs more frequently. Although OALECD relates *melancholy* to both *youyu* and *yiyu*, I would argue that *melancholy* is closer to *youyu* because they are both associated with a temperament or character, lasting long and sometimes can be used in a positive way. However, *youyu* is even more positive because it is a temperament admired and sought by many people. *Shanggan* is similar to *melancholy* in that a *shanggan de* person is also easily affected by emotions such as sadness. S/he tends to feel sad when s/he sees or thinks of something, which often seems incomprehensible to other people. It is quite close to *melancholy* in 'feeling of sadness for no particular reason'.

5.2.8 *Dejected**, *disheartened** & *despondent** and their Chinese equivalents

This section will look at *dejected**, *disheartened**, *despondent** and their Chinese near-equivalents.

5.2.8.1 *Dejected**, *disheartened** & *despondent**

5.2.8.1.1 *Dejected**

Let us start with *dejected*, for which the dictionaries provide the following:

LDCE:

[Date: 1400-1500; Language: Latin; Origin: *dejectus*, from *jacere* 'to throw'] unhappy, disappointed, or sad

OALECD:

adj depressed; sad 沮丧的; 垂头丧气的; 情绪低落的; 郁郁不乐的:

COBUILD:

If you are **dejected**, you feel miserable or unhappy, especially because you have just been disappointed by something. = *despondent*

The dictionaries use 'unhappy', 'sad', 'depressed', 'disappointed', 'miserable' to define *dejected*. OALECD uses four words, three of which are four-character words, to express its meaning: *jusang de*, *tuitousangqi de*, *qingxudiluo de*, *yuyubule de*, in which the last one seems to be far from the meaning core of *dejected* and instead closer to *gloomy*. No entry of *deject* is found in the three dictionaries. The BOE only generates 226 hits of *deject* (v). A simple glance of its concordance reveals that in most cases it is used in the *ed*-form. So I just search for the form of *dejected*, regardless of word class tag. 600 hits of *dejected* are found in the corpus. Its top 50 collocates shows that it colligates with the article-determiners *a* (92), *the* (45), link verbs, predominantly *be* (32), *look* (32), *feel* (11), verbs, like *sit* (4), *stand* (3), *leave* (3), nouns, such as *face* (4), *tone* (3), and prepositions, mainly *about* (6), *over* (3). Its more frequent occurrence with *be* and *look* than *feel* seems to suggest that *dejected* is more extroverted than introverted. In other words, it is more associated with an outward manifestation than an internal state. Its intensifiers cover a wide range of degree and include *very* (12), *utterly* (7), *totally* (6), *rather* (5), *little* (5), *somewhat* (3), *quite* (3).

107 hits of *dejectedly* are found in the corpus and its top 50 collocates show that that the actions which are often done in a *dejected* manner include *say* (12), *sit* (10), *walk* (8), *stand* (4) etc., suggesting that a *dejected* emotion is often manifested outwardly. In its top collocates we do not find any verb that is related to an inner thought or a feeling, which further supports my earlier claim that *dejected* is more characterised by extroversion.

For *dejection*, I find the following in OALECD and COBUILD:

OALECD:

n [U] sad or dejected state; depression 忧郁; 沮丧; 情绪低落

COBUILD:

[N-UNCOUNT] **Dejection** is a feeling of sadness that you get, for example, when you have just been disappointed by something. = *despondency*

As can be seen from the above definitions, the two dictionaries use the same words as for its stem to define or explain *dejection*, like *sadness*, *depression* and *disappointed*, therefore telling us nothing more about its meaning. Turning to its corpus data, 139 hits of *dejection* are found in the BOE. From its top collocates, it can be seen that *dejection* has colligations with prepositions, predominantly *of* (33), *in* (21), *into* (4), modifiers, typically *utter* (5), *deep* (4), and possessive adjectives, like *his* (4), *her* (4). From its word sketch, I cannot find any word that contributes significantly to its meaning due to its low frequency. So I choose not to make a claim about it. This is why I emphasise again and again throughout the thesis that a large corpus is needed to retrieve reliable and useful data.

5.2.8.1.2 *Disheartened**

Now let us move on to look at the *dishearten* family. Among the three dictionaries used for this research, only OALECD includes an entry for *dishearten*:

v [Tn] cause (sb) to lose hope or confidence 使(某人)失去希望或信心; 使(某人)灰心: *Don't let this set-back dishearten you.* 不要因这一挫折而气馁.

However, the other two dictionaries offer an entry for *disheartened*, rather than *dishearten*:

LDCE:

adj formal disappointed, so that you lose hope and the determination to continue doing something

COBUILD:

[ADJ] usu v-link ADJ, oft ADJ by n If you are **disheartened**, you feel disappointed about something and have less confidence or less hope about it than you did before. = discouraged

According to OALECD and LDCE, *disheartened* is characterised by disappointment and losing hope or confidence. However, COBUILD has a mild explanation, i.e. 'disappointed about something and have less confidence or less hope about it than you did it before', thus synonymous with 'discouraged'. Which definition is more adequate?

I turn to the BOE, where all forms of *dishearten* are tagged as a verb. Only 718 hits of *dishearten* (v), including *disheartened* and *disheartening*, are found in the corpus. Its top 50 collocates reveal that *be* (529) tops the list, showing that in most cases *dishearten* is used in the participle form, either *disheartening* or *disheartened*. After inspecting the corpus, I find that *dishearten*, mostly in its participle form, colligates mainly with *get* and *feel*, suggesting that *disheartened* is predominantly inward. From the word sketch for *dishearten* (as a lemma), it can be seen that getting *disheartened* is most frequently caused by *drop* (5), *lack* (3), *defeat* (3), *criticism* (2) and *failure* (2). The collocates

listed in the ‘and/or’ pattern include *frustrate* (4), *disillusion* (2), *discourage* (2), *stun* (2), *confuse* (2), *shock* (2) and *disappoint* (2), suggesting that feeling *disheartened* is triggered by something that has just happened and it is often related to unexpectedness as well. Further, *disheartened* has a broad range of modifiers: *very* (31), *rather* (7), *somewhat* (6), *pretty* (5), *extremely* (4), *especially* (4), *particularly* (4), *terribly* (3), *completely* (3), *totally* (2), *utterly* (2), *deeply* (2) and *profoundly* (2). They can be reorganised according to its degree of losing hope or confidence as follows (in roughly ascending order): *somewhat*, *pretty*, *rather*, *very*, *especially/particularly*, *deeply*, *terribly*, *completely*, *utterly*. We may therefore argue that COBUILD’s definition for *dishearten* ‘disappointed about something and have less confidence or less hope about it than you did it before’ is closer to its meaning.

The only 5 instances of *dishearteningly* from the corpus are all shown below:

<p>Frontier House (Channel 4, Sundays at 8pm). of her passing. Those traces were anyway I think I sensed a great deal more-more conviction of batterers until recently was search for a bone-marrow-transplant donor</p>	<p><i>Dishearteningly</i> <i>dishearteningly</i> <i>dishearteningly</i> <i>dishearteningly</i> <i>dishearteningly</i></p>	<p>, this is an American show (PBS) inspired elusive. After viewing the exterior (as than Papa did-how ineffectual we really low. <p/> There are simple solutions to this difficult, Rob Howard says. So the family</p>
--	--	---

The above lines seem to suggest that *dishearteningly* usually co-occurs with words with a negative connotation, like *elusive*, *low*, *difficult*, *ineffectual*, thus acquiring a negative semantic prosody. It seems that it is still an adverb indicating an emotional manner.

5.2.8.1.3 *Despondent**

This section will explore the *despondent* family. No entry of *despond* is found in the three dictionaries. It only occurs 40 times in the BOE, so we will focus on *despondent*, for which the dictionaries say:

LDCE:

adj [Date: 1600-1700; Language: Latin; Origin: *despondere* 'to give up, lose hope', from *spondere* 'to promise'] extremely unhappy and without hope

OALECD:

adj ~ (about sth) having or showing loss of hope; wretched 失望的; 沮丧的; 消沉的; 苦恼的

COBUILD:

If you are **despondent**, you are very unhappy because you have been experiencing difficulties that you think you will not be able to overcome.

LDCE says that *despondent* originates from the Latin word '*despondere*', meaning 'to give up, lose hope', and implies 'extremely unhappy and without hope'. OALECD claims *despondent* means 'having and showing loss of hope'. Like the definition for *disheartened*, COBUILD's explanation for *disheartened* is also milder than the other two: 'If you are *despondent*, you are very unhappy because you have been experiencing difficulties that you think you will not be able to overcome'.

Tuning to its corpus data (578 hits), its top collocates show that its colligation properties include link verbs, typically *be* (103), *become* (19), *feel* (22), *look* (15), *get* (9), *grow* (5), *seem* (5), and prepositions, particularly *over* (33), *about* (29), *at* (16). Its collocation with *feel*, *look* and *seem* seems to show that *despondent* is both outward and inward. In other words, it is both realised as an emotion involving external manifestations and mentally experienced as an internal emotional state. For its premodifiers, its word sketch shows the following: *increasingly* (12), *irrationally* (2), *thoroughly* (2), *deeply* (2), *totally* (2), *apparently* (2), *somewhat* (2). From the above premodifiers, it can be seen that *despondent* is a serious emotional state which might be undergoing a gradual process since it is frequently modified by *increasingly*, which neither modifies *dejected* nor *disheartened*. Its co-occurring negative feelings in the

'and/or' pattern include *angry* (6), *depressed* (3), *desperate* (3), *frightened* (3) etc. *Angry* seems to suggest that *despondent* is not a totally passive feeling accepting peacefully the unpleasant reality. We will return to this later when we discuss its Chinese equivalents.

The BOE contains 84 occurrences of *despondently* and a randomly selected list

of concordance lines is given below:

<p>earth can I fight that?' Michelle pointed could not remember what it was. He looked favour.' 'I'll co-operate,' Toure replied be dodging the flak, buddy,' Graham said moved to the edge of the clearing and stared was no longer working. So, somewhat getting through them,' Andrei summed up can tell me what to do?' the girl thought 'But I'll never do it,' she told herself in the darkness of the rutted track. <p/> moved, swung, wavered and then dipped, . In a corner the television set blinked and no mistake,' said Sam Gamgee. He stood tide overtakes you. Farewell!' <p/> And so Why had she ruined herself, Trudie thought life.' What's the good of it, he thought hoping it was Bunner, but Lucci's voice said How did last night go?' 'kay,' said Holly eleven, she'd realised he wasn't coming.</p>	<p><i>despondently</i> <i>despondently</i> <i>despondently</i> <i>despondently</i> <i>despondently</i> <i>despondently</i> <i>Despondently</i> <i>despondently</i> <i>despondently</i> <i>despondently</i> <i>despondently</i> <i>despondently</i> <i>despondently</i> <i>despondently</i> <i>despondently</i> <i>despondently</i> <i>despondently</i> <i>despondently</i> <i>Despondently</i></p>	<p>at the television screen, which was showing up at the window, which seemed large and . 'What do you want to know?' Philpott . 'We've blown our one real chance of at the twisted remains of the Audi which , we journeyed to Aberdaron, the nearest . <p/> 'Why not?' Ilmari, the only one who as she exchanged her house-shoes for . 'I think of such mad things to do and he started the car. <p/> His clothing was . The head behind them drooped. without a sound. The blinds were down and with hunched shoulders beside Frodo, and Merry now stood and watched the , as though finally coming out of a stupor , what's the point? Even if I win against , 'Meers did it, Dave. He met the press . 'What's wrong?' demanded Joan. , she switched off the TV, blew out the</p>
---	--	---

From the concordance listing, we can find 'pointed', 'looked', 'replied', 'said', 'stared', 'told', 'stood' etc., showing that *despondently* is still a pure adverb indicating an emotional manner, unlike many emotional adverbs, which have undergone or are undergoing the 'delexicalisation' process.

OALECD defines *despondency* as 'loss of hope or misery' (泄气; 沮丧; 失望; 苦恼). The observation of its corpus data reveals that *despondency* (302 hits) has a colligation with prepositions, particularly *of* (65), *among* (10), *into* (9), *about* (9), *by* (9), *with* (8), and possessive adjectives, like *his* (8), *her* (5). Its word sketch tells us that *despondency* collocates with negative feelings, such as *gloom* (17), *despair* (9), *apathy*

(2), *sadness* (2) and *helplessness* (2). Its frequent co-occurrences with *gloom* and *despair* show that *despondency* is more a feeling focusing on what is going to happen in future rather than an emotional reaction to what has happened.

5.2.8.2 Chinese equivalents of *dejected**, *disheartened** & *despondent**

The BPC contains only 6 instances of *dishearten**, which are all shown below:

- 1) 她只觉得寒冷、**沮丧**(*jusang*)和绝望。She was chilled and ***disheartened*** and desperate.
- 2) 他仍旧愁眉不展，**灰心丧气**(*huixinsangqi*)。He was still gloomy and ***disheartened***.
- 3) 傍晚，当她精疲力竭**垂头丧气**(*chuitousangqi*)地回到家时，她发现杜洛埃来过了。When she arrived at the house at the end of the day, weary and ***disheartened***, she discovered that Drouet had been there.
- 4) 每当她发现他的知识那样丰富，她心中的见解又是那样浅薄的时候，要是同他的像安地斯山一样的智力相比，她就不禁自惭形秽，**心灰意冷**(*xinhuiyileng*)，再也不愿作任何努力了。As such she compared him with herself; and at every discovery of the abundance of his illuminations, and the unmeasurable, Andean altitude of his, she became quite *dejected*, ***disheartened*** from all further effort on her own part whatever.
- 5) 他毁谤亡人，以谰言玷污其美名，用这种下流手段来达到政治上的成功，使有道德之人甚为**沮丧**(*jusang*)。It is ***disheartening*** to virtuous men to see such shameful means resorted to achieve political success as the attacking of the dead in their graves and defiling their honored names with slander.
- 6) 那些经过烟熏火燎的房基是黑糊糊的烟囱（如今叫做谢尔曼的哨兵）令人失望(*ling ren shiwang*)地不断出现。Smoked foundations and the lonesome blackened chimneys, now known as ‘Sherman's Sentinels,’ appeared with ***disheartening*** frequency.

All the Chinese words used to translate *dishearten** are marked in bold: *jusang* (2), *huixinsangqi* (1), *chuitousangqi* (1), *xinhuiyileng* (1), and *ling ren shiwang* (1). We have discussed *jusang* in an earlier section. *Ling ren shiwang* is an interpretative translation, meaning ‘disappointing’. We will skip this word and focus on the other three. 16 hits of *despondent** are found in the parallel corpus. Among them, 5 instances are translated as

(*lingren*) *jusang* and 2 as *chuitousangqi*. 15 occurrences of *deject** are found in the BPC.

Among them, 5 hits are translated as *jusang* and 3 hits as *chuitousangqi*. Taking all of them into consideration, we will examine *chuitousangqi*, *huixinsangqi*, *xinhuiyileng*, *sangqi* and *huixin*. *Huixinsangqi* and *chuitousangqi* both consist of *sangqi*, thus they are quite similar. *Sangqi* and *huixin* are very important concepts which constitute *chuitousangqi*, *huixinsangqi* and *xinhuiyileng*.

5.2.8.2.1 *Sangqi*

Sangqi is glossed as ‘因事情不顺利而情绪低落 [mood is low because things have not proceeded smoothly] downhearted; dejected; crestfallen’ in the XHCBL. It is a verb-object structure formed by *sang* (lose) and *qi* (gas). *Qi* is a Chinese specific notion from the traditional Chinese medicine. It refers to *nengliang* (energy) that exists in the form of gas which flows through the body. If a person has less *qi*, s/he will be less energetic. It has developed many senses, such as spirits/morale/vigour, so *sangqi* means ‘losing vigour or spirits’. The CCL corpus contains 904 occurrences of *sangqi*, out of which 493 hits come from *chuitousangqi* and 170 from *huixinsangqi*. Only in 241 hits it occurs on its own or as a constituent of another phrase, like *ling ren sangqi* (10). *Sangqi* co-occurs with *juewang*/despair (11), *cuozhe*/setback (9), *shibai*/failure (8) and *kunnan*/difficulty (7). It may be argued that the latter three are the main causes for *sangqi*. It seems that *sangqi* is, for the most part, caused by the loss of or less confidence, enthusiasm and determination to go on doing something due to setbacks, failure and difficulties. Hence, it is an emotional reaction to what has happened, similar to *dejected* in this sense. *Sangqi* is found to collocate with *juede*/feel (10), *xiande*/appear

(11), *shenqing*/facial expression (6), *xinli*/in the heart (7). 18 occurrences of *sangqi de yangzi* (appearance) and 9 instances of *shuo* (say) *sangqi hua* (words) are found in the corpus. These all suggest that *sangqi* is characterised by both extroversion and introversion.

5.2.8.2.2 *Chuitousangqi*

Chituousangqi is a phrase formed by two words: *chuitou* (with one's head drooped) + *sangqi*. Therefore, we can claim that it is an expression mainly describing someone's outward manifestation when s/he feels *dejected*. It occurs 492 times in the CCL corpus. Its top collocates include *shuo*/say (42), *yangzi*/appearance (17), *kan*/look (14), *xiande*/appear (9), *shenqing*/ facial expression (6), *yinwei*/because (9) and *bian*/thus (8) etc. *Shuo*, *yangzi*, *kan*, *xiande* all clearly show that *chituousangqi* is predominantly used to describe the external manifestations when s/he is *dejected*. It co-occurrence with *yinwei* (because) and *bian* (thus) shows that *tuichousangqi* has a clear reason for this feeling.

5.2.8.2.3 *Huixin*

Another similar fixed expression is *huixinsangqi*. Before discussing *huixinsangqi*, let us look at *huixin*, which literally means 'grey heart'. In the Chinese culture, 'a grey heart' is conceptualised as the feeling of 'discouraged' or 'losing heart'. *Huixin* occurs 685 times in the CCL corpus and its top collocates are given below:

Table 5.8. 1 Significant collocates of *huixin* in the CCL corpus

Introversion /extroversion	Modifier	Feeling	Others
<i>gandao</i> 35	<i>hen</i> 28	<i>shiwang</i> 45	<i>shi</i> 44
<i>juede</i> 11	<i>wanquan</i> 6	<i>juewang</i> 20	<i>ling</i> 14

<i>qingxu</i> 8	<i>yidian/dian</i> 10/15	<i>qinei</i> 15	<i>rang</i> 12
<i>shuo</i> 33	<i>youxie</i> 5	<i>jusang</i> 8	<i>shibai</i> 15
		<i>shangxin</i> 5	<i>cuozhe</i> 15
		<i>xiaochen</i> 5	<i>kunnan</i> 14
		<i>shiyi</i> 5	<i>yinwei/yin</i> 12/12
			<i>xiwang</i> 8
			<i>fangqi</i> 7

The first column gives the collocates that are related to its extroversion or introversion. It can be seen that introversion dominates the picture, as shown in *gandao/juede* (feel), *qingxu* (mood). *Shuo* (say) also occurs frequently with *huixin*, but it is not a typical verb characterising extroversion in comparison to other apparent extroversion-verbs like *look*, *appear* etc. It is usually preceded by *hen* (very), *wanquan* (totally), *yidian/dian* (a bit), *youxie* (slightly), which shows that it is not a very intense emotional state because its modifiers display various degrees. The feelings it co-occurs with include *shiwang* (disappointed), *juewang* (desperate), *qinei* (discouraged), *jusang* (dejected), *shangxin* (sad), *xiaochen* (low-spirited), *shiyi* (frustrated) etc. Its other significant collocates are *shi/ling/rang* (make), *shibai* (failure), *cuozhe* (setback), *kunnan* (difficulty), *yinwei/yin* (because), *xiwang* (hope), *fangqi* (give up) etc. They seem to suggest that the *huixin* are frequently caused by failure, setback, difficulty and less or loss of hope. The act of giving up might be the result of *huixin*. *Yinwei/yin* appears to show that *huixin* has a clear reason and therefore is quite specific. *Huixin* is an emotional reaction to what has happened, which directly leads to your less or loss of confidence about what you are going to do, so it is a feeling relating the past to the future. It is more inward, as compared with the more outward emotion *sangqi*. It is close to *disheartened* in the above respects.

5.2.8.2.4 *Huixinsangqi*

The CCL corpus generates 170 hits of *huixinsangqi*, which often co-occurs with *shi/make* (17), *ling/cause* (5), *gandao/feel* (12), *shuo/say*(9), *kunnan/difficulty* (7), *cuozhe/setback* (6), *shibai/failure* (5), *qingxu/mood* (5) etc. From its collocates *gandao* and *qingxu*, we can see that, in comparison to *chuitousangqi*, it is more experienced as an inner state. It is also mainly caused by *kunnan*, *cuozhe* and *shibai*.

5.2.8.2.5 *Xinhuiyileng*

The last word I am going to discuss in this section is *xinhuiyileng*, which is formed by *xinhui*, an inverted form of *huixin*, and *yileng* (intention-cold), literally meaning ‘one’s desire is cold’. We can see that sadness in Chinese can also be conceptualised in terms of colour and temperature. The CCL corpus contains 226 occurrences of *xinhuiyileng*, which co-occurs with *shi/make* (24), *rang/make* (11), *ling/make* (8), *xiang/think* (12), *shenghuo/life* (9), *gandao/feel* (7), *xiaochen/low-spirited* (6), *bian/thus* (6), *xiwang/hope* (5) and *youxie/a bit* (5) etc. From its constituent *xinhui* and its collocates *gandao*, *xiang*, *xiaochen*, it can be inferred that it is a word focusing on the inner emotional state. Another salient feature is that, like *xinsui*, it frequently co-occurs with causative constructions, as indicated in its collocates *shi*, *rang*, *ling*.

5.2.8.3 Comparison

We have discussed *dejected**, *despondent** and *disheartened** and their Chinese near-equivalents, i.e. *chuitousangqi*, *huixinsangqi*, *xinhuiyileng*, *sangqi* and *huixin*. We will include *jusang*, which was explored earlier, into the present discussion. Table 5.8.2 summarises their respective features:

Table 5.8. 2 Comparison of *dejected**, *despondent** and *disheartened** and their Chinese equivalents

	<i>dejected*</i> (past) (state) (more outward)	<i>disheartened*</i> (past-future) (state) (more inward)	<i>despondent*</i> (future) (process) (inward-outward)
<i>jusang</i> (past-future) (process) (inward-outward)	√	√	√
<i>huixin</i> (past-future) (state/inward)		√	√
<i>sangqi</i> (past) (state/outward)	√		
<i>xinhuiyileng</i> (past-future) (process/inward)		√	√

It is obvious from the table that *dejected* is more inward and *disheartened* more outward, while *despondent* is both experienced as an inner feeling and realised as an emotion involving physical manifestations. For the Chinese equivalents, both *huixin*, and *xinhuiyileng* focus on the inner emotional state and *sangqi* emphasises the external manifestations, while *jusang* is characterised by both. Among the words in question, *despondent*, *jusang* and *xinhuiyileng* are conceptualised as a gradual process which may change, while all the others are understood as an emotional state which does not change much. *Dejected* focuses on the response to what has just happened and *despondent* describes the emotional state due to the gloomy prediction of what is going to happen in the future, while *disheartened* centres on both. If you are *sangqi*, you are very disappointed about the outcome and *dejected* about it; if you feel *huixin*, you feel *disheartened* about it and have less or lost the confidence and determination to keep doing it; if you feel *xinhuiyileng*, you feel totally *despondent* about it and want to give it up. *Chuituousangqi* and *huixinsangqi* are stronger versions of *sangqi* and *huixin* respectively; they are similar in terms of the extent of losing hope, but the former more highlights the accompanying behaviour, namely ‘drooping one’s head’, while the latter

more focuses on the internal emotional state. *Jusang* covers a wide semantic range; it is both extroverted and introverted; it relates what happened to what is going to happen; it is conceptualized as a dynamic process. Therefore, it corresponds to all three English words.

5.2.9 *Gloom** and their Chinese equivalents

The words to be examined in this section consist of *gloom*, *gloomy*, *gloomily* and their Chinese equivalents---*yinyu* and *menmenbule*. The discussion will start with the examination of the English words.

5.2.9.1 *Gloom**

Let us first look at *gloom*, for which the three English dictionaries provide the following:

LDCE: *n* [singular, U]

① *literary* almost complete darkness ② a feeling of great sadness and lack of hope

OALECD: *n*

① [U] near darkness 昏暗; 阴暗: ② [U, sing] feeling of sadness and hopelessness 忧郁; 忧愁; 失望

COBUILD:

1 [N-SING] the N, oft in/into N

The **gloom** is a state of near darkness.

2 [N-UNCOUNT] also a N

Gloom is a feeling of sadness and lack of hope.

The three dictionaries all agree that *gloom* has two meanings: the original meaning of ‘darkness’ and the metaphorical meaning of a feeling, which is basically described as ‘sadness’, ‘hopelessness’ or ‘lack of hope’. The Chinese equivalents for this feeling sense provided by OALECD are *youyu*, *youchou* and *shiwang*. Actually, *youyu* is the closest in meaning, while *youchou* emphasises worries and *shiwang* is simply disappointment. For a deeper understanding of *gloom*, we need to look at its corpus data.

3439 hits of *gloom* are found in the BOE. From its top 50 collocates, it can be seen that *gloom* colligates predominantly with prepositions, characteristically *of* (526), *in* (149), *for* (91), *over* (67), *into* (43), *about* (33), and the definite article *the* (1219). Its word sketch shows that *gloom* is semantically associated with:

- Object_of: a) *pierce* (22), *lift* (94), *deepen* (20), *lighten* (11), *penetrate* (6), *cast* (16), *encircle* (4), *spread* (8), *gather* (6), *settle* (95), *surround* (4)
b) *predict* (10), *forecast* (5)
c) *dispel* (14), *relieve* (7)
- Subject_of: *descend* (20), *hang* (17), *surround* (12), *deepen* (11), *loom* (10), *envelop* (7), *spread* (7), *lift* (6), *settle* (5), *pervade* (5)

From the collocates in the a) group of the ‘object_of’ pattern, it is clear that *gloom* is conceptualised as something like cloud, fog, gas or light that can be gone into or through and can rise or disperse, go around, move outwards in all directions, come together and remain in the air. The verbs in the b) group are to do with the future and those in the c) group are used in a non-metaphorical way. The subject-of collocates are very much like those in the a) group of the object_of pattern. Although some of the above collocates are used to describe ‘darkness’, in most cases, according to the BOE, *gloom* is used in its metaphorical sense, as illustrated in the following example:

- 1) When the *gloom* finally lifts, the pessimists will be surprised at how much has been going right.

Its word sketch shows that its modifiers include *economic* (46), *deep* (33), *stygian* (18), *global* (9), *grey* (8), *unrelieved* (6), *financial* (5), *dusty* (5), *unrelenting* (4), *oppressive* (4), *pervasive* (4), *heavy* (4) etc. It can be seen that many of them are related to its abstract meaning, like *economic*, *global* and *financial*. *Gloom* is modified by *heavy* to indicate a great degree. The abstract collocates of *gloom* used in the pattern ‘and/or’

include *doom* (452), *despondency* (17), *despair* (17), *pessimism*(5), *depression* (7) etc.

Doom occurs highly frequently due to the fixed phrase ‘doom and *gloom/gloom* and doom’, meaning ‘when there seems to be no hope for the future’. As expected, all the other negative-feeling collocates have to do with ‘lack of hope’.

Let us move on to examine *gloomy*, which is described as the following in the three dictionaries:

LDCE: *adj*

▣▶making you feel that things will not improve = depressing

▣▶sad because you think the situation will not improve = depressed

▣▶dark, especially in a way that makes you feel sad

OALECD: *adj*

▣▶ dark or unlighted, esp in a way that is depressing or frightening 阴暗的, 黑暗的 (尤指使人沮丧或恐惧) ▣▶ (that makes people feel) sad and depressed (使人感到) 忧愁的, 沮丧的

COBUILD:

1 [ADJ] If a place is **gloomy**, it is almost dark so that you cannot see very well.

2 [ADJ] If people are **gloomy**, they are unhappy and have no hope. = despondent

3 [ADJ] If a situation is **gloomy**, it does not give you much hope of success or happiness. = grim

All the dictionaries divide the meaning of *gloomy* as a feeling into two categories: one is ‘sad’ and the other is ‘making sad’. OALECD lists it as one meaning and only puts ‘that makes people feel’ in the brackets. The Chinese equivalents it provides include *youchou de* and *jusang de*. The former focuses on ‘worries’ and the latter on ‘frustration’. For a deep understanding of its meaning, we might as well look at its corpus data. 2824 hits of *gloomy* are found in the BOE. Its top collocates show that its colligational properties include the article-determiners *a* (470), *the* (430) and the preposition *about* (64). Its word sketch reveals that it semantically prefers *picture* (93),

forecaster(72), *prognosis* (23), *prediction* (68), *outlook* (61), *forecast* (43), *view* (43), *prospect* (22), *scenario* (16), suggesting that *gloomy* is more hopeless about one's prospect or future than the outcome of a specific event. This is also reflected in the collocates of the 'pp_about_p' pattern: *prospect* (15) and *future* (7). It seems that *gloomy* is more outward than inward as we can find *look* (73), *sound* (12), *seem* (11) and *appear* (5) in the collocates of the pattern 'adj_comp_of'. *Feel* only occurs 22 times in this pattern.

For *gloomily* (382), let us look at its word sketch directly. Its 'verb' pattern tells us that *gloomily* are semantically associated with verbs which can be done in an emotional manner, typically *say* (65), *stare* (20), *think* (10), *sit* (6), *reply* (6) and *nod* (6), out of which only *think* is an introverted verb, so we may argue that *gloomily* is more outward than inward. For its adverb premodifier, I only find *rather* (9). The verb it frequently premodifies is *predict* (11).

5.2.9.2 Chinese equivalents of *gloom**

Now let us move on to look at the Chinese equivalents for the *gloom* family. *Gloom** occurs 151 times in the BPC and its top Chinese equivalents are shown below:

Table 5.9. 1 Chinese equivalents of *gloom** in the BPC

Feeling	Original meaning
<i>youyu</i> : 11	<i>yin'an</i> :11
<i>yinyu</i> : 10	<i>you'an</i> ; 10
<i>menmenbule</i> : 7	<i>hei'an</i> : 7
<i>yinchen</i> : 8	<i>an</i> : 7
<i>you</i> (<i>chou/lv</i>): 7	

We will focus on the expressions in the first column as the words in the second column are related to its original meaning of 'darkness'. Since *youyu* has been discussed in an

earlier section and *yinchen* is a word describing countenance rather than expressing feelings directly, so we will only discuss *yinyu* and *menmenbule*. 7 hits of *gloom** are translated as *you* (*chou/lü*), showing that *gloom* is linked with the Chinese concept of *you*, which means worry or concern. It is not the focus of this research, so we will exclude it from our discussion. The following example seems to show that *gloom* is a more serious state than *depression*:

- 2) 不断地将他过去的处境和现在的处境相对比,表明平衡正向坏的一面倾斜,于是产生了一种终日忧郁(*youyu*)或者至少是消沉(*xiaochen*)的心态。Constant comparison between his old state and his new showed a balance for the worse, which produced a constant state of *gloom* or, at least, *depression*.

5.2.9.2.1 *Yinyu*

The significant collocates of *yiyu* (667 hits) are shown below:

Table 5.9. 2 Significant collocates of *yinyu* in the CCL corpus

Extroversion	Modifier	Feeling	Others
<i>lianse</i> 27	<i>hen</i> 13	<i>chenzhong</i> 10	<i>tianqi</i> 8
<i>shenqing/biaoqing</i> 27/16	<i>shifen</i> 8	<i>beiai</i> 5	<i>qifen</i> 21
<i>shense</i> 13	<i>feichang</i> 7	<i>jusang</i> 5	<i>longzhao</i> 13
<i>kan</i> 26	<i>dai</i> 15	<i>tongku</i> 5	<i>chenmo</i> 13
<i>xiande</i> 19			<i>chensi</i> 12
<i>muguang/yanshen</i> 18/17			<i>chenmen</i> 7
<i>yanjing/yan</i> 16/7			<i>xingge</i> 5
<i>lianshang/lian</i> 13/11			<i>xinqing</i> 16
<i>miankong</i> 9			<i>qingxu</i> 14
<i>shengyin</i> 10			

The first column shows the collocations which are related to physical manifestations, either via expressions or voice, including *lianse* (countenance), *shenqing/biaoqing/shense* (facial expression), *kan* (look), *muguang/yanshen* (expression in one's eyes), *xiande* (*appear*), *yanjing/yan* (eyes), *lianshang/lian/miankong* (on one's face/face), *shengyin* (voice). It seems to suggest that *yinyu* is more experienced as an emotion with physical manifestations than an inner feeling. Its modifiers include *hen/shifen/feichang*

(very) and *dai* (slightly). It also co-occurs with other sad feeling words, such as *beiai* (sorrowful), *jusang* (dejected) or *tongku* (anguished). In XHCBL, *yinyu* is glossed as:

1. (天气)低沉郁闷; (气氛)不活跃 (of weather) gloomy; dismal; (of atmosphere) not lively
2. 忧郁, 不开朗 depressed; closed and gloomy

This explains why we find *tianqi* (weather), *qifen* (atmosphere), and *longzhao* (shroud) in the last column. The frequent expression would be some place is ‘*longzhao zai yinyu de qifen zhong*’ (lit. *yinyu* hangs over), where *longzhao* is a typical Chinese word to describe cloud, fog or shadow. Therefore it is similar to *gloomy* in terms of collocating with words describing cloud or fog. This is because emotions and feelings are often understood metaphorically in terms of various forms of substance – gaseous, fluid, and solid (see Kövecses, 2000, Chapter 5). *Chenmo* (silence), *chensi* (contemplation), and *chenmen* (dreariness) are often associated with *yinyu*, suggesting that *yinyu* has a quality of quietness. It is also used to talk about a person’s character, so we can say *yinyu* could be extended in time. Another feature is that it frequently co-occurs with *xinqing* (frame of mind) or *qingxu* (mood), but seldom with *ganqing* or *qinggan*, which is more believed to be the Chinese counterpart of *feeling*. This feature is also shared by most of the Chinese sadness words discussed. It probably shows that in the Chinese culture sadness is more understood as a *xinqing* or *qingxu*, rather than a *ganqing* or *qinggan*; it also confirms the idea suggested in R. Zhang & Ooi (2008) that *ganqing* and *qinggan* are not exact equivalents of *feeling*.

5.2.9.2.2 *Menmenbule*

The CCL corpus contains 435 hits of *menmenbule* and its top collocates are shown below:

Table 5.9.3 Significant collocates of *menmenbule* in the CCL corpus

Extroversion	Introversion	Others
<i>shuo/shuodao</i> 51/7	<i>xinli/xinzhong</i> 14/6	<i>ziji</i> 13
<i>zou</i> 10	<i>juede</i> 10	<i>bian</i> 10 <i>qilai</i> 9
<i>xiande</i> 7	<i>gandao</i> 8	<i>yizhi</i> 17 <i>zhengtian</i> 11
<i>lian</i> 5	<i>xinqing</i> 7	<i>zongshi</i> 10 <i>changchang</i> 7
<i>biaoqing</i> 5	<i>qingxu</i> 5	<i>weishenme</i> 11
	<i>xiang</i> 26	<i>yinwei</i> 6

From the above table, it can be seen that *menmenbule* is frequently used to describe an outward action which is done in an unhappy manner, like *shuo/shuodao* (say) or *zou* (leave). We also find *xiande* (appear), *lian* (face), *biaoqing* (facial expression) in the first column, which all point to its characteristic of extroversion. However, we can also find *xinli/xinzhong* (in the mind), *juede/gandao* (feel), *xinqing* (frame of mind), *qingxu* (mood) and *xiang* (think), which are closely related to its feature of introversion. Therefore, *menmenbule* is a combined feeling of both extroversion and introversion. The collocates in the last column give us an idea about its other features. *Ziji* (oneself) suggests that it seems to be a feeling confined to oneself. *Bian* (change) and *qilai* (used after a verb to indicate the beginning of an action) appear to show that *menmenbule* starts at a certain point of time. *Yizhi/zongshi* (all the time/always), *changchang* (often), and *zhengtian* (all the day) seem to show it often lasts for a rather long time. *Weishenme* and *yinwei* (because) tell us that usually such a feeling has a clearly-identified reason. It is post-modified by *hen* (very/10). Another semantically similar four-character idiom is *yuyuguahuan*, which is more formal and literary. The original meaning of *men* in Chinese is ‘tightly closed or sealed’. If someone bottles up all his/her unhappy feelings or worries in the mind and does not lessen them by sharing them with others, s/he will get *menmenbule*.

5.2.9.3 Comparison

In this section we will compare *gloom** with *youyu*, *yinyu* and *menmenbule*. *Youyu* is often semantically associated with positive words, such as *xiao* (smile), *meili* (beautiful), *langman* (romantic), *wenrou* (gentle) and *wangzi* (prince); *menmenbule* semantically prefers *zhongshi/yizhi* (always/all the time); *yinyu* is more associated with words indicating or related to outward manifestations; *gloomy* has a totally different set of collocates which are associated with the future, such as *prospect*, *future*, *outlook*, *forecast* etc.

None of these three Chinese equivalents corresponds exactly to *gloomy*. *Youyu* is more an inward feeling which is extended in time rather than a momentary occurrence and has no specific reason. However, sometimes it is also viewed as something very beautiful. *Yinyu* is more manifested outwardly and could last short or long. Both *youyu* and *yinyu* can be related to a person's character. *Youyu* is closer to *melancholy*, though they differ in many aspects. *Yinyu* is similar to *gloomy* in that they are both outwardly manifested. However, *gloomy* can be used to modify a situation that makes people sad. *Menmenbule* is more specific because it has a definite reason. In comparison with *youyu*, it lasts shorter. It will fade away with the gradual disappearance of the cause. It is a combination of extroversion and inversion. *Menmenbule* is the least intense in the Chinese *gloomy* group. If people are *gloomy*, they think the situation will not improve. *Menmenbule* is a feeling when things are not going as wished. If people are *menmenbule*, they are not satisfied with what happened and they are thinking about the past. When people are *gloomy*, they are thinking about the future.

5.2.10 *Upset** and their Chinese equivalents

5.2.10.1 *Upset**

This section will investigate *upset*, for which the following is found in the three dictionaries:

LDCE:

- adj ① [not before noun] unhappy and worried because something unpleasant or disappointing has happened ② **be upset with sb**: if you are upset with someone, you are angry and annoyed with them ③ **upset stomach**: an illness that affects the stomach and makes you feel sick
- v [T] [Date: 1700-1800; Origin: Probably from upset 'to set up, raise' (15-17 centuries)] ① **【MAKE SOMEBODY UNHAPPY】** to make someone feel unhappy or worried ② **【CHANGE SOMETHING】** to change a plan or situation in a way that causes problems ③ **【MAKE SOMETHING FALL】** to push something over without intending to ④ **【DEFEAT】** to defeat an opponent who is considered to be much better than you
- n ① [U and C] worry and unhappiness caused by an unexpected problem ② [C] when a person or team defeats an opponent who is considered to be much better than them ③ **stomach upset**: an illness that affects the stomach and makes you feel sick

OALECD:

- v ① [I, Tn] (cause sth to) become overturned or spilt, esp accidentally 打翻, 弄翻 (某物) (尤指无意地) ② [Tn] disrupt (a plan, etc) 打乱, 扰乱 (计画等) ③ [Tn] (a) distress the mind or feelings of (sb) 使 (某人) 苦恼或心烦: (b) cause (sb) to feel ill by disturbing his digestion 使 (某人) 感到肠胃不适
- n ① [U, C] upsetting or being upset 翻倒; 扰乱; 不安 ② [C] stomach disorder 肠胃不适 ③ (in sport) unexpected result (运动) 出乎意料的结果或成绩.

COBUILD:

1 [ADJ] usu v-link ADJ, oft ADJ by/about n If you are **upset**, you are unhappy or disappointed because something unpleasant has happened to you.

[N-COUNT] **Upset** is also a noun.

2 [VERB] V n, V pron-refl If something **upsets** you, it makes you feel worried or unhappy.

3 [VERB] V n If events **upset** something such as a procedure or a state of affairs, they cause it to go wrong.

[N-COUNT] **Upset** is also a noun.

4 [VERB] V n If you **upset** an object, you accidentally knock or push it over so that it scatters over a large area.

5 [N-COUNT] supp N

A stomach **upset** is a slight illness in your stomach caused by an infection or by something that you have eaten.

[ADJ] ADJ n **Upset** is also an adjective.

From the emotion-related definitions in LDCE and COBUILD, we can see that the determinants of *upset* are worry, unhappiness and disappointment. OALECD uses *distress*, a more difficult word, to define *upset*, and thus gives no valuable information. LDCE's explanation of the phrase 'be *upset* with somebody' as 'be angry and annoyed with somebody' seems to introduce anger into its composition. Now let us look at the corpus data. 3866 hits of *upset* (n) are found in the BOE and its top collocates reveal that *upset* as a noun, as expected, colligates with the articles *the* (342), *an* (582), and prepositions, typically *of* (454), *by* (81), *at* (72), *over* (65). According to its word sketch, its significant adjectival modifiers include *big* (243), *major* (150), *stunning* (58), *emotional* (45), *huge* (37), *first-round* (28), *minor* (16), *shocking* (14), *potential* (13), and its nominal modifiers are predominantly *stomach* (125) and *tummy* (28). Interestingly, *upset* is often modified by *stunning* or *shocking*, showing that *upset* is much to do with unexpectedness. This confirms that disappointment is a main factor of getting *upset*. From the collocates above, it can also be seen that *upset* is conceptualised as something that can change in size with intensity, as indicated in *big*, *huge*, *major* and *minor*. Despite the big difference in frequency between *emotional* and *stomach/tummy*, there is no obvious inclination that *upset* without a modifier might refer to an emotional or stomach *upset* because both occur quite frequently. Contexts have to be relied on to decide what it is. The words used in the pattern of 'and/or' include *diarrhoea* (10), *rash*

(5), *nausea* (4), *headache* (6), *cold* (6), *streak* (8), *controversy* (6), *anger* (5), *pain* (7), *trouble* (4). Only the last four words have to do with an emotional *upset*.

Now let us look at *upset* as an adjective. 4070 hits are found in the corpus and its top collocates tell us that *upset* (adj.) colligates strongly with prepositions, particularly *about* (221), *with* (106), *by* (83). According to its word sketch, it is predominantly associated with link verbs, typically *be* (570), *get* (448), *become* (101), *feel* (75), *look* (56), *seem* (47), *sound* (25) and *appear*(14), suggesting that it is both extroverted and introverted, because *sound*, *appear*, *seem* and *look* are closely related to outward manifestations. It frequently collocates with amplifiers, predominantly *very* (191), *really* (64), *extremely* (27), *terribly* (19), *deeply* (14), emphasisers, mainly *particularly* (11), *genuinely* (10), and downtoners, typically *pretty* (20), *quite* (22) and *little* (4). However, we can see that it collocates much more frequently with amplifiers than downtoners.

Its word sketch show that the emotion words it co-occurs with in the ‘and/or’ pattern include *angry* (153), *distressed* (14), *confused* (14), *tired* (17), *nervous* (18), *sad* (14), *unhappy* (10), *anxious* (9), *mad* (9), *stunning* (9), *frightened* (9) etc. From the above collocates, it is clear that *angry* is the most frequent collocate of *upset*, co-occurring with it 153 times in this pattern, much higher than other negative emotion words. This seems to suggest that being *upset* is closely related to *anger* and they are two co-existent feelings stimulated by unpleasant and unexpected events.

17124 hits of *upset* as a verb are found in the BOE and its top 50 collocates show that *upset* as a verb strongly colligates with *be* (4335), and prepositions, typically

to (1664), *by* (1146), *about* (1009), *at* (631), *with* (512), *over* (224). This is because a) it is often used in passive voice and b) in many instances a verb tag is mistakenly assigned to an adjective. Its word sketch reveals that its frequent objects include *balance* (216), *applecart/apple-cart/cart* (64), *stomach* (34), *seed* (38), *resident* (31), *everyone/everybody* (24), *rhythm* (17), *neighbour* (15), suggesting its metaphorical usage is more common. Its subjects include *something* (29), *news* (21), *death* (19), *decision* (19), *incident* (16), *comment* (14), *criticism* (12), *attack* (12), *remark* (11), *teenager* (10), *fact* (10), *event* (10). It seems that most of them, if not all, are very much likely to be connected to emotions. Some of them have obvious negative inclinations themselves, such as *death*, *criticism*, *incident*, *death*, *violence*, *attack* and *risk*, which are prone to *upset* people, while others may need to have negative premodifiers to make people sad, like *remark*, *comment*, *news*, *decision*, *fact*, *event* etc. Its predominant intensifiers include amplifiers, such as *very* (1044), *really* (468), *deeply* (122), *extremely* (102), *greatly* (25), emphasisers, like *particularly* (108), *genuinely* (27), and the downtoner *mildly* (5). A simple glance of their frequencies will show that *upset* (v) is much more likely to be intensified by amplifiers than downtoners. Its word sketch shows that the words used in the ‘and/or’ pattern are closely related to emotions, such as *shock* (61), *hurt* (49), *disappoint* (32), *worry* (27), *offend* (19), *depress* (14), *frustrate* (18), *surprise* (18), *annoy* (17), *embarrass* (16), *anger* (15), *frighten* (14), *disturb* (11), *concern* (11). These negative emotion words reveal that an *upset* feeling might be mixed with one or more of these shown feelings, especially those highly frequent ones, like *shock*, *hurt*, *disappoint*, *worry*. The fourth sense of *hurt* in LDCE is glossed as ‘insult somebody: to

make someone feel very *upset*, unhappy, sad etc'. The following example from LDCE illustrates well the *worry* component of *upset*:

- 1) If you are the victim of a burglary, the emotional *upset* can affect you for a long time.

One might be *unexpected* about what has happened, and get *sad* and *disappointed*.

Based on what happened so unexpectedly, the experiencer might get *worried* about the future and get *angry* about the cause to make this happen. This might be the most typical case for *upset*. It covers the whole process. So far, we seem to be able to conclude that *upset* is a mixed feeling of *unexpectedness*, *unhappiness*, *disappointment*, *worry*, and *anger*. In certain situations, some factors might be less significant and some may get dominant.

5.2.10.2 Chinese equivalents of *upset**

OALECD defines *upset* as 'distress the mind or feelings of (sb) (使 (某人) 苦恼或心烦)', where the Chinese equivalent concepts used to translate *upset* are *kunao* (worried) and *xinfan* (perturbed). Let us look at the parallel corpus data. *Upset** occurs 60 times in the corpus and its top 4 Chinese equivalents are *fannao* (5), *nanguo* (5), *shengqi* (4) and *zhaoji* (3). The most frequent Chinese translations of *upset* include *fannao* (worries) and *nanguo* (sad), *shengqi* (annoyed) and *zhaoji* (anxious/worried). *Fannao* is the closest in meaning to the Chinese equivalents provided by OALECD—*kunao* and *xinfan*. However, the parallel corpus tells us that the other three are also its equivalents. Actually, none of these is the exact or close equivalent to the English *upset*. It seems that no single Chinese word qualifies as its equivalent, because all of the above-mentioned elements

constitute the feeling of *upset*. However, in translating a certain text segment, the translator only focuses on one of the aspects based on the context and uses a semantically similar word to convey the intended meaning prominent in that context.

Consider the following examples:

- 2) ‘Well, almost everything,’ Frank amended hastily, disturbed by the expressions on their faces. He tried to look cheerful, for he did not believe in *upsetting* ladies. ‘唔，差不多全烧光了，’ 弗兰克显然对她们脸上的表情感到有点为难，才连忙纠正说。他想要显得愉快一些，因为他不主张叫小姐太太们烦恼(*fannaο*)。
- 3) The blush was not difficult for she was breathless and her heart was beating like a drum. ‘Rhett, I’m so sorry about what I--I said to you that night--you know--at Rough and Ready. I was--oh, so very frightened and *upset* and you were so--so--’ She looked down and saw his brown hand tighten over hers. ‘And--I thought then that I’d never, never forgive you! But when Aunt Pitty told me yesterday that you--that they might hang you--it came over me of a sudden and I--I--’. She looked up into his eyes with one swift imploring glance and in it she put an agony of heartbreak. ‘Oh, Rhett, I’d die if they hanged you! 她脸上的红晕是不难做到的，因为她已经喘不过起来，心也似敲鼓般的怦怦直跳。’ 瑞德，我很抱歉，我对你---我那天晚上对你说的---你知道---在拉无雷迪。那时我---啊，我多么害怕和着急 (*zhaoji*)，而你又是那么---那么---’她眼睛朝下，看见他那只褐色的手把她的手腕抓得更紧了。’所以---那时我想我永远永远也不饶恕你！可是昨天皮蒂姑妈突然告诉我说，你---说他们可能会绞死你---这真把我吓倒了，所以我---我---’她抬起头来，用急切祈求的目光注视着他的眼睛，她的目光中还含着揪心的痛苦。‘啊，瑞德，要是他们把你绞死了，我也不想活了！’
- 4) ‘What will Mother say when she hears?’ He looked up in sudden anguished apprehension. ‘You wouldn’t be telling your mother a word and *upsetting* her, now would you?’ ‘妈听到了会怎么说呢?’ 他忽然惊慌失措地抬起头来。‘你总不至于向你妈透露让她难过(*nanguo*)吧，会吗?’
- 5) She picked up the cologne bottle and took a large mouthful, carefully rinsed her mouth and then spit into the slop jar. She rustled down the stairs toward the two who still stood in the hall, for Pittypat had been too *upset* by Scarlett’s action to ask Rhett to sit down. He was decorously clad in black, his linen frilly and starched, and his manner was all that custom demanded from an old friend paying a call of sympathy on one bereaved. In fact, it was so perfect that it verged on the burlesque, though Pittypat did not see it. He was properly apologetic for disturbing

Scarlett and regretted that in his rush of closing up business before leaving town he had been unable to be present at the funeral. 她拿起香水瓶，往嘴里倒了一大口，漱了半天，吐在了痰盂里。她赶紧下了楼，看见他们还在过厅里站着，朝他们二人走去，皮蒂姑妈正为思嘉举动而生气(*shengqi*)，没顾上请瑞德坐下。瑞德郑重其事地穿着一身黑衣服，衬衫上镶着褶边，而且是浆过的，一切举止也都符合一位老朋友向失去亲人的人表示慰问的样子，一切都是那么周到，甚至到了可笑的地步，但皮蒂姑妈并没有察觉，他这么晚前来打搅，一本正经地向思嘉表示了歉意。

In Example (2), Frank modified his previous remark by saying ‘almost everything’ and tried to look cheerful in order not to make ladies worried. The translation focuses on the constituent meaning part of *fannaο* (worry). In Example (3), When Scarlett heard that Rhett would be hanged, she got frightened and anxious. *Zhaoji* (anxiety) was emphasised in this context. Example (4) shows that he did not want to make his mother *nanguo* (sad) by hiding the secret. The situation for Example (5) was a bit complicated. Scarlett drank a lot of brandy on the funeral day of her husband and her behaviour was unacceptable at that time. Aunt Pitty was so *shengqi* (annoyed) that she forgot to ask Rhett to take a seat when he visited her to show his sorrow.

It can be seen that different words denoting quite distinct Chinese concepts are used to convey the meaning of *upset* in these translations. It seems to support the claim that the congruence between languages is a more textually valid type of equivalence (González, Mackenzie, & Álvarez, 2008). Seen in isolation, *upset* is an emotion word which expresses a mixture of several feelings, i.e. *sadness*, *worry*, *anxiety* and *annoyance*. Chinese does not encode so many components into one single emotional word. It seems that Chinese emotional concepts tend to be more discrete, with each emphasizing one emotional aspect.

5.2.11 *Bitter**, *agony**, *anguish** and *tongku*

In English, *anguish* and *agony* are synonymous, both meaning ‘extreme mental or physical suffering or pain’. *Bitter* is something different. However, they are all connected to the feeling of *tongku* in Chinese. To find out how they differ from each other and why they are all related to *tongku*, we need to examine them in the corpus. Let us look at *bitter* first.

5.2.11.1 *Bitter**

Below are the entries for *bitter* as an adjective from the three dictionaries:

LDCE:

☐ feeling angry, jealous, and *upset* because you think you have been treated unfairly ☐ [only before noun] making you feel very unhappy and *upset* ☐ a bitter argument, battle etc is one in which people oppose or criticise each other with strong feelings of hate and anger ☐ having a sharp strong taste like black coffee without sugar ☐ unpleasantly cold

OALECD:

☐ having a sharp taste like aspirin or unsweetened coffee; not sweet 苦的; 有苦味的 ☐ difficult to accept; causing sorrow; unwelcome 难以接受的; 引起悲伤的; 不受欢迎的 ☐ caused by, feeling or showing envy, hatred or disappointment ☐ piercingly cold 严寒

COBUILD:

- 1 [ADJ] In a **bitter** argument or conflict, people argue very angrily or fight very fiercely.
- 2 [ADJ] If someone is **bitter** after a disappointing experience or after being treated unfairly, they continue to feel angry about it.
- 3 [ADJ] usu ADJ n A **bitter** experience makes you feel very disappointed. You can also use **bitter** to emphasise feelings of disappointment.
- 4 [ADJ] **Bitter** weather, or a **bitter** wind, is extremely cold.
- 5 [ADJ] A **bitter** taste is sharp, not sweet, and often slightly unpleasant.

From the above definitions, it can be seen that the original meaning of *bitter* as a bad taste has been extended and used in different domains and developed different meanings.

My analysis will focus on the senses which have to do with feelings. What is shared by

these three dictionaries is that it denotes an unpleasant feeling or it is used to modify something which can cause such a feeling. LDCE defines *bitter* as ‘feeling angry, jealous, and *upset* because you think you have been treated unfairly’, and ‘making you feel very unhappy and *upset*’. OALECD says that *bitter* is ‘difficult to accept; causing sorrow; unwelcome’ and ‘caused by, feeling or showing envy, hatred or disappointment’. COBUILD claims that ‘If someone is *bitter* after a disappointing experience or after being treated unfairly, they continue to feel angry about it’. To summarise, *bitter* as a feeling is associated with ‘angry/resentful’, ‘jealous/envious’, ‘*upset*’, and ‘disappointed’.

Now let us turn to look at how *bitter* behaves in the BOE, which yields 13274 hits of *bitter* as an adjective. An inspection of its top 50 collocates reveals that *bitter* colligates with the articles *a* (3027), *the* (2043), prepositions, predominantly *about* (307), *with* (189), link verbs, typically *be* (813), *feel* (124), and nouns it modifies. It can be clearly seen from its collocates that *bitter* is more often used in its metaphorical sense: *rival* (435), *dispute* (259), *blow* (226), *disappointment* (201), *experience* (198), *battle* (177), *memory* (163), *row* (146), *feud* (139), *enemy* (135), *war* (120), *rivalry* (95), *attack* (103), *fight* (100), *debate* (92), *divorce* (85), *fighting* (82), *division* (82), *conflict* (78), *irony* (76), *struggle* (76), *argument* (68) etc. *Taste* only occurs 272 times. *End* (264) and *pill* (205) occur frequently due to the fixed phrases *to the bitter end* and *a bitter pill*. In the sense of ‘piercing cold’, it collocates frequently with words related to coldness, such as *cold* (222), *wind* (65) and *winter* (63). In the sense of ‘caused by, feeling or showing envy, hatred or disappointment’ (LDCE), it has a semantic

preference for things involving opposition, antagonism or hostility, such as *rival/rivalry, dispute, debate, conflict, fight/fighting, feud, struggle, attack, blow, row, war, division, divorce, argument, enemy, irony*. It also semantically prefers for things causing people unhappy and *upset*, typically *experience, memory, disappointment*. Its word sketch shows that its adverbial modifiers include *extremely (19), increasingly (125), particularly (42), rather (18), quite (13), slightly (30), little (22), somewhat (10)* etc. It also shows that *bitter* often co-occurs with *resentful (40)* and *angry (104)*, as exemplified in the following examples:

- 1) It is vital for the bereaved person to express their grief, for grief cannot be suppressed. The feelings must go somewhere. Sometimes unexpressed grief may turn inwards, festering into a ***bitter*** and resentful outlook on life, or it may manifest as chronic physical tension.
- 2) Labour groups have consulted only organizations from their own constituency, treating these as the representatives of a genuine `community`, and in doing so ignoring representatives of local business or excluding the often ***bitter*** and resentful local white working class. This is a strategy for subverting dialogue while appearing to encourage it and is not part of the process of open corporatism advocated here.
- 3) Either the system of justice has not operated properly and you've been unjustly punished, or it has operated properly and you've got what you deserve. If you choose the first you become angry, ***bitter*** and resentful. Our society is full of angry, bitter and resentful people. Some of them appear in the media, voicing their complaints about women priests, or homosexuals, or the youth of today ...
- 4) Mr Cooper said a secret survey of prison staff revealed they were severely stressed and their morale was devastatingly low. `Across the board, the findings reveal staff are angry, ***bitter***, sunk in cynicism and apathy, isolated, ignored, badly treated seriously stressed. It is simply appalling that prisons staff are so shabbily treated.' But Queensland Corrective Services Commission director-general Keith Hamburger said the situation was not as bad as indicated. Moves had been made to reduce stress on staff in the State's prisons.
- 5) We saw how Helen (Chapter 2) died angry and ***bitter*** at the man who jilted her, only to ruminate continually on this theme in her novels today.

Bitter is associated with *outlook* in Example 1, *working class* in Example 2, *people* in Example 3, and *cynicism and apathy* in Example 4. It seems to suggest that *bitter* could be a long-term feeling or an attitude toward life. Example 5 shows a typical case of *bitter* in which the feeling, like angry, is more aiming at someone else, such as the man who jilted her, rather than oneself.

Turning to *bitterly* (3715 hits), its word sketch²⁰ shows that, roughly speaking, it semantically prefers three groups of words: one is to do with opposition, typically verbs, such as *complain* (280), *divide* (161), *oppose* (182), *contest* (104), *fight* (77), *resent* (73), *criticise/criticize* (58), *dispute* (26), *comment* (17), *quarrel* (14), *argue* (16), *retort* (5), etc. Some verbs may not be prototypical opposing verbs, like *comment*, but they still assume an opposing side. The second one is related to verbs showing or feeling sad emotions, such as *regret* (87), *weep* (37), *sob* (14), *remark* (10), *smile* (23), *laugh* (35), *cry* (17). The third group include adjectives denoting unpleasant feelings, such as *disappointed* (426), *disappointing* (63), *unhappy* (15), *angry* (10), *resentful* (6), showing opposing feelings, such as *ironic* (12), *hostile* (11), *critical* (18), or relating to coldness, like *cold* (331). In its sense associated with a feeling, it seems that if it involves an opposition, it tends to be related to hatred or anger; if it describes a passive feeling one suffers mentally, it is likely to be an unacceptable and unwelcome feeling, more like ‘*upset*’ and ‘*unhappy*’.

²⁰ Some of the frequency figures presented here are the sum of two figures from its two patterns: ‘v’ and ‘pre-modifier v’. There is no pointing in distinguishing between its positions in my research, so I group them together manually.

Let us move on to look at *bitterness*, which is the nominal form of *bitter*, mainly in the sense of ‘difficult to accept; causing sorrow; unwelcome’ and ‘caused by, feeling or showing envy, hatred or disappointment’. The corpus yields 2684 hits of *bitterness*. Its top 50 collocates and its concordance show that it has a strong colligation with prepositions, typically *of* (604), *with* (94), *toward(s)* (71), *between* (42), *over* (54), *about* (44), *at* (53), *without* (38), *against* (21). The preposition *toward(s)* confirms my earlier claim that *bitter* is often a quite active feeling, unfriendly aiming at someone else. *Bitter* also colligates with verbs, typically *express* (20), *cause* (14), *create* (12), *feel* (10), modifiers, like *great* (26), *deep* (7) and quantifiers, typically *much* (27), *little* (11). Its word sketch reveals that it frequently co-occurs with negative feelings, such as *anger* (117), *resentment*(55), *hatred* (42), *frustration* (26), *despair* (20), *rage* (14), *jealousy* (13), *envy* (10), *hostility* (8) etc.

Now let us turn to its Chinese counterparts. 218 instances of *bitter** occur in the BPC. Its original meaning as a *bitter* taste occurs only 13 times. Another clearly discernible translation equivalent is *leng* (cold) (12). It is interesting that *bitter/bitterly* is often used to indicate a degree and its Chinese translations for this sense include a wide range of adjectives (18) and adverbs (25, including 13 instances of *hen*). The translations of its other metaphorical uses are summarised in the table below:

Table 5.11.1 Chinese equivalents of *bitter** in the BPC

Feeling	Others
<i>tongku</i> : 44	<i>ku</i> : <i>ku/kunan/shouku/kunku</i> :10
<i>xinsuan</i> : 9	<i>jianke/ keke/kebo</i> :8
<i>xiao:kuxiao</i> 8+ <i>lengxiao</i> 1	<i>nanguo/jiannan/zao/laosao/</i>
<i>kunao/fannaο/aonao</i> :6	<i>yuanshengzaidao</i> : 7
<i>nanshou/nanguo/beishang/beitong</i> : 6	others:42

As was mentioned earlier, *bitter* has been extended into many domains and developed dozens of meanings, which basically can be classified into two categories, i.e. ‘unpleasant’ or ‘causing an unpleasant feeling’. Chinese has many words to describe such situations, so it can be translated as many words which convey such meanings. From the above table, we can see that the most frequent translation is *tongku* (44). It is also associated with other feelings. In 9 instances it is translated as *xinsuan* (heart-sour). In 9 instances it is used to modify laugh/smile, where the Chinese translations are *kuxiao* (bitter-laugh) and *lengxiao* (cold-laugh), as exemplified in the following example (*bitter* is translated as *lengxiao*):

- 6) 如果从构成这一群人中的每一个男人、每一个女人和每一个尖嗓门的孩子的口中爆发出轰笑，海丝特·白兰或许可以对他们所有的人报以桀傲的冷笑(*lengxiao*)。... Had a roar of laughter burst from the multitude--each man, each woman, each little shrill-voiced child, contributing their individual parts--Hester Prynne might have repaid them all with a *bitter* and disdainful *smile*.

6 of the translations are close to ‘annoyed and worried’, such as *fannao*, *kunao*, *aonao*, each emphasizing a different aspect. 5 of them are to do with sadness, such as *nanshou*, *nanguo*, *beishang* and *beitong*. For the non-feeling translations, 10 of them are associated with *ku* (bitterness) in its metaphorical sense, such as *kunan* (distress), *shouku* (suffering), *kunku* (hardship). This is because, in Chinese, the taste of *bitter* is also conceived of as something unpleasant or unacceptable, like hardship, suffering and pain. 8 of them are connected with *ke* (acrimonious) like *jianke*, *keke* and *kebo*. *Ke* roughly means ‘harsh, unpleasantly stern or unkind’, often linked with envy or hatred. 5 of them are associated with a difficult life, including *nanguo* (hard-live), *jiannan* (hard), *zao* (bad), *laosao* (complaint) and *yuanshengzaidao* (complaints are heard everywhere).

From the above analysis, it can be seen that the taste of *bitter* in English can be mapped onto many domains, resulting in various metaphorical meanings. Some of them are shared by Chinese, such as *kunan* (bitter-difficulty), *shouku* (suffer-bitterness), *kunku* (difficulty-bitterness). It seems that the closest Chinese equivalent for *bitter* in the sense of feeling is *tongku*, with the taste of *bitterness* mapped onto the domain of feeling, as exemplified in the following examples:

- 7) 然后，他又恢复了冷静、痛苦(*tongku*)的感觉。
Then he came down to cold, *bitter* sense again.
- 8) 他抬起头来这个冷酷、痛苦(*tongku*)的现实，使他猛地清醒了。He raised his head, and the cold, *bitter* reality jarred him into wakefulness.
- 9) 在他最需要的时候失去她，使他加倍痛苦(*tongku*)。
His loss seemed all the more *bitter* now that he needed her most.

In the following example, *nanshou* was used to translate *bitter* because *tongku* is a relatively strong word to denote a feeling of pain:

- 10) 就像大人們的煩惱也是煩惱一樣，他忘記煩惱並不是因為他的煩惱不怎麼沉重和難受(*nanshou*)，而是因為一種新的、更強烈的興趣暫時壓倒並驅散了他心中的煩惱--就像大人們在新奇感受的興奮之時，也會暫時忘卻自己的不幸一樣。 Not because his troubles were one whit less heavy and *bitter* to him than a man's are to a man, but because a new and powerful interest bore them down and drove them out of his mind for the time -- just as men's misfortunes are forgotten in the excitement of new enterprises.

5.2.11.2 *Agony* & anguish**

In this section, I will compare and contrast *agony** and *anguish**. Look at their dictionary entries:

***agony*:**

LDCE:

n plural agonies [U and C] [Date: 1300-1400; Language: Late Latin;
Origin: *agonia*, from *Greek*, 'trouble, great anxiety', from *agon*
'competition for a prize'] **▶**very severe pain **▶**a very sad, difficult, or
unpleasant experience

OALECD: [U, C]

extreme mental or physical suffering (精神或肉体的) 极大痛苦

COBUILD:

[N-UNCOUNT] also N in pl **Agony** is great physical or mental pain. = torment

anguish:

LDCE:

n [U] *written* [Date: 1100-1200; Language: Old French; Origin: *angoisse*, from *Latin* *angustiae* 'extreme upset', from *angustus* 'narrow']
mental or physical suffering caused by extreme pain or worry >**anguished** *adj*

OALECD:

n [U] severe physical or mental pain (肉体的或精神的) 极度痛苦: >
anguished *adj* feeling or expressing *anguish* 感到痛苦的; 表现痛苦的

COBUILD:

[N-UNCOUNT] **Anguish** is great mental suffering or physical pain.

(WRITTEN) *Mark looked at him in anguish.*

From the definitions shown above, we can see that *agony* has one basic meaning, namely 'great/extreme mental or physical suffering/pain'. LDCE breaks it down into two senses: 'very severe pain' and 'a very sad, difficult, or unpleasant experience'. According to the dictionary definitions, *anguish* is extremely similar to *agony* except that *anguish* is 'caused by extreme pain or worry' (LDCE).

To find what they exactly mean and how they differ, we need to look at the corpus data. For the sake of space, I will not offer a detailed analysis, but only summarise the relevant research findings. First, their colligation properties are summed up as follows:

agony (5299):

preposition: *in* (1188), *of* (1112), *over* (56)

definite article: *the* (1166)

adjectival possessive: *his* (110), *our* (45), *her* (58), *their* (41), *own* (16)

anguish (2592):

preposition: *of* (619), *in* (245), *over* (77)

definite article: *the* (495)

adjectival possessive: *his* (100), *her* (87), *their* (62), *my* (20), *its* (13)

It can be seen that their colligational patterns are almost the same. Turning to their sketch difference, their common patterns can be summarised as follows (*agony/anguish*, only significant collocates are listed):

Object_of: *suffer* (103/27), *express* (5/18), *experience* (8/14), *cause* (12/36), *feel* (16/27), *ease* (11/12)

And/or: *ecstasy* (54/5), *despair* (5/16), *grief* (4/19), *pain* (37/102), *frustration* (4/13), *anger* (5/22), *fear* (7/20)

A_modifier: *mental* (20/92), *sheer* (18/5), *terrible* (18/8), *emotional* (8/14), *physical* (12/4), *deep* (3/14), *personal* (11/17), *private* (7/14), *great* (26/34)

Considering their occurrences (5299/2592) in the corpus, *anguish* co-occurs much more frequently with other negative feelings, such as *despair*, *grief*, *pain*, *frustration*, *anger* and *fear*. What is also remarkable from the above listing is that *anguish* occurs much more frequently with *mental* (35‰) and much less frequently with *physical* than *agony* (3.7‰). This difference might suggest that *anguish* without any modifier tends to denote a physical pain. Its more frequent co-occurrence with *emotional* than *agony* (5.4‰/1.5‰) also support this claim. For *agony*, its collocating frequencies with *mental/physical* (20/12) do not show a striking difference, so it can be argued that *agony* is quite neutral in isolation. Both tend to be modified only by amplifiers, such as *sheer*, *terrible*, *deep*, rather than downtoners. The *anguish-only* patterns show that its possessor is all in the family circle i.e. *family* (19), *mother* (7), *father* (3), probably because it is often caused by worry, and that it is also often associated with other negative feelings, like *fury* (8), *isolation* (7), *heartache* (5), *loneliness* (4), *hurt* (3), *desperation* (3). The *agony-only* patterns reveal that it is semantically associated with *prolong* (44), *endure* (34), *inflict* (12), *shorten* (3) etc., which might imply that *agony*,

unlike *anguish*, is more conceptualised as a dynamic process rather than a static state. Its unique adjectival modifiers are mainly *excruciating* (12) and *absolute* (31), suggesting that *agony* tends to be more intensified by amplifiers. In addition, it is often used in the pattern ‘agony of X’. Its concordance shows that in this pattern an enormously wide range of words can occur as a cause of *agony*: *failure, hunger, heat, homesickness/seasickness, helplessness, expectation, impatience, grief, guilt, shame, remorse, repentance, loss, separation, jealousy, fear, regret, worry, poverty, unhappiness, embarrassment, fright, indecision, irresolution, uncertainty* etc. This seems to suggest that in terms of cause *agony* covers a broad range in the scale of seriousness. An inspection of its concordance reveals that it occurs much more frequently with mental causes than physical causes in the BOE, which might be because we talk about our mental experiences much more than physical ones. Consider the following examples:

- 11) His voice was hurried, nervous. 'For the first time in my life I wished I smoked, ' he said, looking down at his feet in an *agony* of embarrassment. 'I like you, Trudie, I like you a lot . ' I 'm no good . ' He shook his head . 'What 's good when it 's at home ? ' 'Respectable !'
- 12) The cat sat in front of the bird cage in an agony of *frustration* at being so near and yet so far. 猫无可奈何(*wuke'naihe*)地坐在鸟笼前 , 眼看着鸟儿近在咫尺 , 可怎么也够不着。(http://dict.cn/frustration)

In Example 11, embarrassment is the cause of *agony* or the mental pain itself. But in Chinese, in such a situation, the experience of getting embarrassed is never associated with *tongku*. Similarly, in the case of Example 12, the failure to reach the bird is far from being *tongku*. So we may argue that *agony* is not necessarily intense or severe.

Now let us move on to look at *agonize/agonise*. Before the corpus analysis, the relevant dictionary definitions are examined:

LDCE

to think about a difficult decision very carefully and with a lot of effort

~ **over/about**

OALECD:

v [I, Ipr] ~ (**about/over sth**) suffer great anxiety or worry intensely

(about sth) (对某事物) 极度忧虑或担心

> **agonized, -ised** *adj* expressing agony 表示痛苦的 **agonizing, -ising**

adj causing agony 使人痛苦的

COBUILD

[VERB] V over/about n If you **agonize over** something, you feel very anxious about it and spend a long time thinking about it.

agonizing/agonising:

1 [ADJ] Something that is **agonizing** causes you to feel great physical or mental pain. ● **agonizingly** [ADV]

2 [ADJ] **Agonizing** decisions and choices are very difficult to make.

agonized/agonised:

[ADJ] usu ADJ n **Agonized** describes something that you say or do when you are in great physical or mental pain

A close inspection of the above dictionary entries reveals that *agonise/agonize* (*over/about*), as a verb, means quite differently from its nominal form. It has a pretty specific meaning, closely related to anxiety, worry and long-time thinking. However, OALECD's explanations of its *-ing* and *-ed* forms, i.e. 'causing agony' and 'expressing agony', show that these two forms are still close in meaning to its nominal form. COBUILD's paraphrases of them confirm this hypothesis.

The corpus can tell us whether this is true. The BOE generates 1195 hits of *agonize* and 1399 hits of *agonise*. I examine them together for their top 50 collocates. As expected, *agonise/agonize* strongly colligates with *over* (489) and *about* (92). Its collocates show that the possible reasons to *agonize/agonise* include *wait* (93), *death* (64), *decision* (60), *pain* (46), *choice* (20), *question* (17), *dilemma* (8), *defeat* (9),

slowness (12), *loss* (11) etc. Out of the 2594 instances in total, 1753 hits are in the *-ing* form and 612 in the *-ed* form, leaving only 229 hits of other forms. The collocates of *agonizing/agonising* suggest that it semantically prefers words indicating a time period, predominantly *day* (33), *moment* (26), *hour* (17), *minute* (18), *time* (17), *month* (16), *week* (16), *year* (13), and *night* (12), revealing that to *agonise/agonize* is a dynamic *process* (14), which could be as short as a moment or as long as a year. Its other nominal collocates include *wait* (89), *death* (62), *decision* (58), *pain* (44), *choice* (20), *question* (16), *slowness* (12), *detail* (12), *loss* (11), *defeat* (9), *dilemma* (8), *experience* (8) and *reappraisal* (8). Among them, *decision*, *choice*, *dilemma* and *reappraisal* are concerned with judgment or evaluation. Other collocates are all associated with physical or mental pain.

Now let us move on to examine *agonized/agonized* and *anguished* together. The relevant corpus findings are summarised as follows:

anguished (849): *cry* (48), *look* (17), *voice* (11), *tone* (7), *expression* (6),
eye (6), *moan* (4), *wail* (4), *howl* (3), *scream* (3)
agonized/agonised (612): *scream* (10), *cry* (9), *face* (8), *expression* (6),
sound (6), *look* (6)

It can be seen that both are semantically associated with words to do with visible or audible manifestations, such as *cry*, *scream*, *look*, *voice*, *expression*, *sound*. However, it appears that, compared with *agonized/agonized*, *anguished* is more frequently associated with physical pain, as indicated in its collocates, such as *cry*, *scream*, *moan*, *wail*, *howl* etc., which apparently denote physical pain. From the discussion so far, it is clear that *agonized/agonised* and *agonizing/agonising* diverge considerably in meaning from their stem *agonize/agonise*, but go towards their root.

Turning to their Chinese equivalents, 88 hits of *agony** are found in the BPC. A summary of the search findings is given below:

Table 5.11. 2 Chinese equivalents of *agony** in the BPC

Feeling	<i>tongku</i> 44, <i>zhemo</i> 3, <i>jutong</i> 3, <i>nanweiqing</i> 2, <i>nanshou</i> 2, <i>tongchu</i> 2
Idiom	<i>zhuanlan</i> 5
Left out	2
Others	25
Total	88

In half of the instances, *agony** is translated as *tongku*. Its other translations include *zhemo* (a stronger version of *tongku*), *jutong* (lit. severe physical pain), *tongchu* (another formal word for pain), *nanshou* (a less formal word for pain), *nanweiqing* (from *agony of embarrassment*), *zhuanlan* (from *agony column*) etc. In still many other cases, it is translated in various ways because it has a wider semantic space, with causes ranging from very minor to very serious matters.

For *anguish*, the BPC only yields 46 hits. The search result is shown below:

Table 5.11. 3 Chinese equivalents of *anguish** in the BPC

Feeling	<i>tongku</i> 21 (<i>jidu tongku</i> 5), <i>tongchu</i> 1, <i>zhemo</i> 3, <i>xinzhong ...jutong/xinli...jitong</i> 5, <i>beitong/fenwai nanguo</i> 4, <i>kunao/kumen</i> 5
Physical pain	3
Others	4
Total	46

It can be clearly seen that although *anguish* can mean both physical and mental pain, it is only used in its physical sense 3 times in the corpus. This might be attributed to the composition of the parallel corpus since fiction talks much more about feelings than physical pains. Among the feeling group, 21 instances are translated as *tongku*, including 5 instances of *jidu tongku* (extremely *tongku*); one instance is translated as

tongchu, which is highly similar in meaning to *tongku* but more formal; in 3 of them, *zhemo*, which roughly means ‘torture’, are used. 5 of them are translated by using split constructions ‘xinzhong jutong’ and ‘xinli...jitong’, which are understood in terms of metaphorical images of what is physically happening in one’s heart. What they depict is more or less the same as *tongku* does, but this way of description exhibits a more vivid picture. *Anguish* is also associated with *grief*, because *beitong* or *fenwai* (extremely) *nanguo* is used in 4 translations. Consider the following example:

- 13) 这种不正常比邦妮的死还要严重，因为邦妮死后初期的悲痛 (*beitong*) 现在已逐渐减轻，她觉得那个惨重的损失可以默默地忍受了。 This wrongness went even deeper than Bonnie's death, for now the first unbearable ***anguish*** was fading into resigned acceptance of her loss.

5 of the translations are associated with *kunao/kumen* (vexation/gloom). Obviously, *anguish* focuses more on the pain or suffering, regardless of reason. This is quite different from the Chinese *tongku*, which is more specific. So from the above analysis, it seems that *anguish*, defined as ‘great mental or physical pain or suffering’, consists of similar semantic components to *tongku*. However, it covers a wider range of pain than *tongku* because, as noted earlier, it is also associated with *grief*, *vexation*, *gloom* or *worry*.

5.2.11.3 *Tongku*

This section will discuss the Chinese counterpart shared by this group of words based on the BPC, i.e. *tongku*. For *bitter*, as was presented earlier, OALECD does not include *tongku* in its Chinese explanations. What it offers for *anguish* and *agony* are both *jidu tongku* (extremely+*tongku*). It seems that, according to this dictionary, no English word is equivalent to *tongku* itself. XHCBL defines *tongku* as follows:

身体或精神感到非常难受 : pain; suffering; agony; physical suffering or mental pain: ~ 的生活 painful life | 得了这种病 , 非常 ~ 。 The disease causes great pain and suffering.

Interestingly, it lists neither *anguish* nor *bitter* here. Let us look at how it behaves in the CCL corpus, where 17259 hits of *tongku* occurs and its significant collocates are shown in Table 5.11.4 (p. 241):

As shown in Table 5.11.4, its collocates are classified into five categories. The collocates in the first column are relevant to the characteristic of extroversion: *lian/lianshang* (face/on the face), *biaoxian* (manifest), *biaoqing* (facial expression), *kan/kandao* (look/see), *yanjing* (eyes), *yangzi* (appearance). *Gandao/ganjue/ganshou* (feel), *neixin* (inner heart), *xinzhong/xinli* (in the heart) and *xin* (heart) point to its attribute of introversion. So *tongku* could be a facial expression (*biaoqing*), frame of mind (*xinqing*) or feeling (*ganqing/ganshou*). The verbs in the second column show that *tongku* has semantic preferences for words of causation, typically *rang/ling/shi/jiao* (cause/make), *zaocheng/dailai/dai/chansheng* (bring about/lead to), words related to a state (*chongman*/be filled with, *xianru*/get caught into), suffering (*zaoshou/shou*), enduring (*renshou/ren/chengshou/jingshou*), relieving (*jiechu/ jietuo/baituo*) and lessening (*jianqing/jianshao*). These verbs basically cover the whole process of *tongku*. The third column lists its classifiers, typically *zhong* (kind) in *zhezhong* (this kind) and *nazhong* (that kind), and modifiers, such as *da/juda/jida* (great/ enormous), *nanyirensou/bukan* (unbearable), *hen/feichang/shifen* (very), *shen* (deep), *wanfen/ji/jidu/ji* (extremely) and *yidian* (a little). It seems that *tongku*, in most cases, is intensified by amplifiers and only in a small number of cases by downtoners.

Table 5.11. 4 Significant collocates of *tongku* in the CCL corpus

Introversion/Extroversion	Verb	Classifier & Modifier	Noun		Feeling
<i>lianshang/lian</i> 161/109	<i>shi/ling</i> 787/167	<i>zhezong</i> 423	<i>jingshen</i> 299	<i>shibai</i> 73	<i>ai</i> 168
<i>biaoxian</i> 83	<i>rang/jiao</i> 326/119	<i>nazhong</i> 132	<i>linghun</i> 74	<i>maodun</i> 140	<i>xingfu</i> 152
<i>biaoqing</i> 164	<i>dailai/dai</i> 341/160	<i>yiqie/suoyou</i> 216/81	<i>xinling</i> 93	<i>jiannan</i> 81	<i>xiwang</i> 153
<i>kan/kandao</i> 213/89	<i>zaocheng</i> 178	<i>xuduo</i> 145	<i>sixiang</i> 73	<i>kunnan</i> 63	<i>huanle</i> 268
<i>yanjing</i> 85	<i>chansheng</i> 61	<i>da/juda/jida</i> 426/153/84	<i>guocheng</i> 162	<i>bing/jibing</i> 74/72	<i>kuaile</i> 166
<i>yangzi</i> 71	<i>chongman</i> 185	<i>shen</i> 111	<i>jingli</i> 294	<i>zainan</i> 71	<i>beishang</i> 114
	<i>xianru</i> 137	<i>hen/feichang</i> 663/210	<i>tiyan</i> 62	<i>zhemo</i> 264	<i>beiai</i> 101
<i>gandao</i> 684	<i>renshou/ren</i> 368/89	<i>shifen</i> 180	<i>jingyan</i> 62	<i>zhengzha</i> 169	<i>juewang</i> 136
<i>juede</i> 148	<i>chengshou</i> 128	<i>bukan</i> 224	<i>shenghuo</i> 312	<i>jian'ao</i> 72	<i>fannaoyang</i> 109
<i>ganjue</i> 144	<i>zaoshou</i> 165	<i>nanyirensou</i> 69	<i>rensheng</i> 119	<i>si/siwang/siqu</i>	<i>shiwang</i> 92
<i>ganshou</i> 97	<i>jingshou</i> 73	<i>wanfen</i> 121	<i>jiating</i> 63	222/123/67	<i>kongju</i> 90
<i>neixin</i> 320	<i>jiechu</i> 215	<i>ji/jidu</i> 101/103	<i>aiqing</i> 82	<i>buxing</i> 129	<i>fennu</i> 84
<i>xinzhong/xinli</i> 184/127	<i>jietuo/baituo</i> 89/75	<i>yidian</i> 93	<i>shiqing</i> 96	<i>shiqu</i> 115	<i>bu'an</i> 78
<i>xin</i> 157	<i>jianqing</i> 224		<i>xianshi</i> 75	<i>routi /shenti</i> 140/73	<i>gudu</i> 70
<i>xinqing</i> 98	<i>jianshao</i> 85		<i>shehui</i> 76	<i>shenyin</i> 140	<i>zhizhong</i> 196
<i>ganqing</i> 87			<i>lishi</i> 59		<i>zhong</i> 992
			<i>shengming</i> 129		<i>ziji</i> 620
			<i>renlei</i> 117		<i>ci</i> 142
			<i>huiyi/jiyi</i> 152/56		

Like English, Chinese uses the same word *tongku* to express both physical and mental pain. However, only three collocates in the table are clearly associated with its physical sense: *routi/shenti* (body) and *shenyin* (groan); in contrast, we find a great number of occurrences of *jingshen/xinling* (mind), *linghun* (soul) and *sixiang* (thought). *Tongku* is a *jingli/tiyan/jingyan* (experience) or a *guocheng* (process), which composes *shenghuo/rensheng/shengming* (life). As *renlei* (human kind), we have our *tongku*, which might be associated with *huiyi/jiyi* (memory), *jiating* (family), *shehui* (society), *shiqing* (event), *xianshi* (reality) or *lishi* (history). It may be caused by *shibai* (failure), *maodun* (dilemma), *si/siwang/siqu* (die/death), *buxing* (misfortune), *shiqu* (loss), *jiannan* (hardship), *kunnan* (difficulty), *bing/jibing* (disease) or *zainan* (disaster). Suffering *tongku* is a *zhemo* (torment) or *jian'ao* (torture). Some people may *zhengzha* (struggle) to get out of this situation. *Ai* (love) and *xiwang* (hope) may mitigate *tongku*. Both *kuaile* (joy) and *huanle* (gaiety) are the antonyms of *tongku*. A life with less *tongku* is *xingfu* (happy) *de*. *Tongku* is a close companion of *juewang* (despair), *fannaoy* (vexation), *beishang/beiai* (sorrow), *shiwang* (disappointment), *kongju* (fear), *fennu* (indignation), *gudu* (loneliness) or *bu'an* (uneasiness). People often have mixed feelings of *tongku* and other negative feelings. If someone is said to be *tongku* and *juewang*, we can know that s/he is extremely *tongku*. If someone is *tongku* and *gudu*, we may see that s/he is not very *tongku*. *Tongku* highly frequently occurs in the pattern 'tongku zhizhong/tongku zhong' (in *tongku*), which is similar to *agony* or *anguish*. *Tongku* is often a feeling confined to

oneself (*ziji*), suggesting that *tongku* is often something one has to digest on one's own.

5.2.11.4 Comparison

Now let us compare all these words. First, we will look at how *tongku* and *bitter* differ. *Bitter* as a feeling is a metaphorical extension of a taste, so it can only be used to express an inner emotional state. Unlike *tongku*, it does not have a semantic preference for words indicating outward manifestations, like (facial) *expressions*, *face*, *show*, *eyes* etc. However, like *tongku*, it often collocates with experience or memories, which can cause a *bitter* feeling. *Bitter* can be a passive emotional reaction towards something unacceptable, unwelcome or sorrow-causing, but if it is concerned with an opposition, i.e. 'I feel bitter towards someone else', it tends to be an active feeling mixed with *hatred*, *anger*, *rage*, *jealousy* or *hostility*. *Tongku* is a combination of metaphors, i.e. *tong* (pain) and *ku* (bitter). It has developed into a word denoting *tong*, both mental and physical, and the mental *ku*, with its original meaning of physical *ku* lost. It is mainly used to describe an internal state, but it is also semantically associated with outward manifestations. Both *bitterness* and *tongku* are conceptualised as something which is changeable in size, depth and amount etc., so both can be preceded by modifiers such as *deep*, *great*, *much*, *little* etc. Both semantically prefer *feel*, *express*, *cause*, and other negative-feeling words, like *despair*, *disappointment*. What differs between them is that *tongku* is a good companion of *fannaο* (vexation), *beishang* (sorrow), *beiai* (sorrow) *gudu* (loneliness) and *fennu* (indignation). Except *fennu*, all the other feelings are quite passive. *Fennu* may grow

out of *tongku* in certain circumstances, but people are much more inclined to *shou/reushou/chengshou* (endure) *tongku*. In contrast, *bitterness* is more associated with *anger, resentment, jealousy, hatred, rage, envy, hostility*, which might be caused by the more frequent use of *bitter* in the sense of ‘I feel bitter towards someone else’. If used in this sense, the sufferer is quite active because s/he is unsatisfied with the present situation and trying to get rid of it by turning the fire to other people. So, *bitter* is equivalent to *tongku* only in the non-opposition sense.

Now let us come to the comparison between *anguish* and *tongku*. Unlike *bitterness*, *anguish* implies extreme pain or suffering and can hardly be modified by downtoners, while *tongku* could be a less intense feeling and sometimes intensified by downtoners, such as *youxie* (slightly), *xiangdang* (fairly). *Anguish* and *tongku* both have a semantic preference for verbs like *suffer, feel, express, experience, cause, ease* etc. They are both semantically associated with negative feelings, such as *despair, frustration, fear* and the like. *Anguish* has a semantic preference for family words, such as *family, father, mother*, suggesting that *anguish* is more associated with family. *Tongku* does not have such an inclination.

Agony often occurs in the pattern ‘the *agony* of X’, where X indicates the cause or accompanying behaviour, including *failure, hunger, heat, homesickness/seasickness, helplessness, expectation, impatience, grief, guilt, shame, remorse, repentance, loss, separation, jealousy, fear, regret, worry, poverty, unhappiness, embarrassment, fright, indecision, irresolution, uncertainty* etc. In terms of cause *agony* covers a much broader range in the scale of seriousness. However,

tongku hardly co-occurs with *impatience*, *grief*, *shame*, *remorse*, *repentance*, *fright*, *fear*, *worry*, *poverty*, *indecision*, *irresolution*, *uncertainty*, or *embarrassment*. Compared with *agony*, *tongku* is a more specific emotion, emphasising the unacceptability of what happened to the sufferer and the incompetence to solve the problem or change the situation. It is a quite passive emotional state.

It can be seen that, from the above analysis, *tongku* is only partly equivalent to the three words, overlapping with each in certain aspects. *Bitter* indicates a less intense feeling than *tongku*, only denoting the inner emotional state. *Anguish* and *agony* are both extroverted and introverted, with the former more conceptualised as a static emotional state and the latter more understood as a dynamic process that could be *endured*, *prolonged*, *shortened*. In terms of the spectrum of causes, *agony* has the broadest range of coverage, followed by *anguish* and *bitterness*. *Anguish* and *agony* are translated as *tongku* only if they fall within its scope. Take ‘agony of embarrassment’ for example, it is an unpleasant feeling or mental pain in the general sense, but it is not what the Chinese *tongku* covers.

Tongku is a more introverted and more specific emotion. It is perceived as both a process and a state. It fulfils most of the functions of *anguish*, *agony* and *bitter* in its non-opposition sense, but it has a narrower semantic coverage than both *agony* and *anguish*. The typical case for *tongku* is that someone is suffering and not able to get out of it due to the incompetence to untie the mental knot. Just as the pattern ‘xianru...*tongku*’ conveys, *tongku* is like a pit one is likely to fall into but is difficult to get out of. One may feel helpless or desperate before they free themselves from the

trap. But people who have fallen in the trap are quite passive in the sense that they tend to accept what has happened to them and, in most cases, do not intend to take any action.

5.3 Metaphors in sadness expressions in Chinese and English

A systematic corpus-based comparison of sadness expressions between English and Chinese has been made in the previous section. Before embarking on an overall comparison of their commonalities and differences, it is necessary to talk about metaphors as most of the examined items are metaphorical conceptualisations. Metaphor was regarded as merely a literary device and hence outside the scope of linguistics. It was brought to public attention in linguistics by Lakoff & Johnson's pioneering *Metaphors We Live By* (1980), in which it is characterised as a conceptual phenomenon whereby a target domain is structured and understood with reference to a more basic source domain. It is in fact pervasive in human understanding. Based on Lakoff & Johnson's *Metaphors We Live By*, I summarise the metaphorical mappings between source domains and target domains of all the metaphorical sadness expressions examined in this research in Table 5.13.1 (p. 247).

From the inventory of the expressions shown in Table 5.13.1, we can find that Chinese seems to have more metaphorical sadness expressions at the lexical level (including prefabricated expressions) than English. This is due partially to the fact that Chinese words are composed of characters. If a word contains a character which is a metaphor, the whole word becomes a metaphor as well, like *(bei)tong*,

Table 5.13.1 Comparison of metaphors in sadness expressions in Chinese and English

Chinese metaphorical sadness expression	Source domain	Target domain (feeling)	English metaphorical sadness expression	Source domain	Target domain (feeling)
<i>(bei)tong</i> /pain	physical injury (ontological)	grief			
<i>tongku</i> / painful and bitter	physical pain+taste (ontological)	anguish	<i>bitter</i>	taste (ontological)	unhappy
<i>(bei)shang</i> /hurt	physical injury (ontological)	sorrow			
<i>shangxin</i> /heart is hurt	physical injury (ontological)	sad	<i>brokenhearted</i>	physical injury (ontological)	extremely sad
<i>xinsui</i> / heart is fragmented	physical injury (ontological)	extremely sad			
<i>huixin</i> /grey heart	colour (ontological)	disheartened			
<i>xinhuiyileng</i> /grey heart+cold intension	colour+temperature (ontological)	despondent	<i>deject</i> /cause to fall down	spatial orientation	despondent
<i>(you)yu</i> /accumulate in a closed container (liver)	container	melancholy			
<i>yayi</i> /press down+ press down	spatial orientation	depressed	<i>depress</i> /press down	spatial orientation	very unhappy
<i>yinyu</i> /cloudy+ accumulate in a closed container (liver)	weather (ontological) + container	melancholy	<i>gloomy</i>	darkness (ontological)	sad & hopeless
<i>iyu</i> /press down+accumulate in a closed container (liver)	spatial orientation +container	depression			
<i>bu gaoxing</i> /not+high+mood	spatial orientation	unhappy	<i>upset</i> /overturn	spatial orientation	unhappy, annoyed, anxious & worried
<i>menmen(bule)</i> /tightly closed in a closed container (heart)	container	gloomy			
<i>nanguo</i> /hard+live	live a life (structural)	sad			
<i>qingxu diluo</i> /mood +low	spatial orientation	depressed			

where only *tong* is a metaphor.

What can also be seen from the table is that they are quite similar in terms of metaphorical mappings as both Chinese and English map the domains of physical injury, taste, container and spatial orientation onto the domain of feeling. Both cultures conceptualise mental experience (mental pain or feeling) in terms of physical experience (physical pain, injury or taste) and think of and talk about sadness as DOWN and happiness as UP. For example, in Chinese, there are *qingxu diluo* (mood+down), *yayi* (press down), *yiyu* (press down + (of sorrow, anger etc.) pent-up), *gaoxing* (high+mood), as shown in Table 5.13.1. In English, we can find *depress* (press down), *deject* (cause to fall down), *upset* (overturn), as shown in Table 5.13.1, and *low/high-spirited* which are not discussed and therefore not shown in the table.

According to Lakoff & Johnson (1980), the mapping of attributes in the domain of spatial orientation onto other conceptual domains is referred to as orientational metaphors. They claim that these spatial orientations are based on the nature of the human body and the way the body operates in our physical environment. Lakoff & Johnson (1980) observe that HAPPY/GOOD IS UP and BAD/SAD IS DOWN in English linguistic expressions. They argue that human beings walk erect because they have the ability to overcome the gravitational force, so the erect body is associated with a positive connotation and the prostrate posture with a negative one. This is the reason why UP has positive connotations and DOWN has negative ones, which seems to be shared by Chinese (Lan, 1999, 2002). The fact that Chinese and English cultures

conceive of UP and DOWN in a similar way may be attributed to common human experience, among which walking erect is considered as the essential one.

When we shift the attention from the node word (the examined items) to the wider context and look at the collocations shown in earlier sections, we find that both Chinese and English cultures metaphorically conceptualise emotions in terms of gaseous or fluid, so the image of ‘being full’ in a container is associated with a strong feeling. In English, if someone is filled with an emotion, s/he feels it very strongly. We can say someone is filled with *admiration/joy/happiness/horror/fear/anger/doubt/remorse* etc. This is similar to the description of Chinese emotions. *Chongman* (be filled with) in Chinese is a typical verb used to describe an intense emotion, such as *youyu, beitong, beishang, beiai, shanggan, tongku*. It is interesting to note that in both cultures, metaphorical expressions are mainly used to describe strong feelings, while a modifier or quantifier is used instead to describe a less intense or weak feeling.

It is not surprising that there are similarities in metaphorical conceptualisations of emotions between these two cultures because metaphor is grounded in embodied experience, an idea that has been dominant in cognitive linguistic metaphor research from the very beginning (Lakoff & Johnson, 1980, 1999). We construe our emotions in terms of our physical activities, so our emotional experience is closely related to our physical experience. We are all human beings and share a lot in common in human physical experience, which leads to similar metaphorical conceptualisations in the emotion domain. In fact, ‘most of our normal conceptual system is metaphorically structured’; that is, ‘most concepts are partially understood in terms of other concepts’

(Lakoff & Johnson, 1980: 56). Metaphor is a cognitive tool for conceptualising our world and it plays a central role in the construction of social reality; metaphor is rooted in cultures and pervasive in human languages. Actually, not only metaphors, but the whole human conceptual system emerges from everyday human experience (Sweetser, 1990: 1). This partially explains why all human languages share some similarities.

On the other hand, Table 5.13.1 shows that Chinese and English also display some differences in metaphorical sadness expressions. Firstly, as mentioned earlier, some Chinese emotion words, like *beishang*, *beitong*, *youyu*, *menmenbule* shown in Table 5.13.1, are composed of two parts and only one of them is metaphor (the Chinese characters in the brackets are not metaphors). Obviously, this does not apply to English. Secondly, both cultures have their own specific ways of perceiving sadness. The Chinese culture conceives of sadness in terms of weather, colour, temperature and living a life, such as *nanguo* (hard +live), *xinhuiyileng* (grey heart+cold intension), *yin(yu)* (cloudy). Chinese has a specific sadness concept *yu*, which has developed its meanings in the order of ‘fanmao’ (luxuriant), ‘zuzhi/bise’ (block), ‘youchou/youyu’ (sorrow and worry), ‘yuanhen’ (resentment), ‘baonu’ (rage) and the like (*Hanyu Da Cidian*: 858). In Modern Chinese, it is mainly used in disyllabic words, meaning 1) ‘pent-up sorrow’, as in *yinyu*, *yiyu*, 2) ‘be held up and accumulate’, as in *(you)yu*, *yujie*, *yuji*, and 3) ‘luxuriant’, as in *yuyucongcong*. In the modern discourse, the meaning of ‘pent-up sorrow’ remains active, but the meanings of ‘yuanhen’ and ‘baonu’ have become rare. Its meaning of sadness is associated with a blocked image: the sad feeling in the form of *qi* is held up and accumulates gradually in one’s body. *Ganqi yujie* (liver+qi+block) is a

typical expression to describe the situation when someone's sad feeling has been held up and accumulates in one's liver. The traditional Chinese medicine believes that if *qi* is held up in the liver, it will lead to great sadness, i.e. *yu*. Another Chinese specific concept is *men* (tightly closed), a less serious emotional state, as in *menmenbule*. *Men* occurs when someone is thinking about something unhappy and does not share it with others. It seems to be understood in terms of a closed container, but it is specifically associated with the image of the heart. The typical expression is *menzai xinli* (tightly closed in the heart), because the heart is believed to be the locus of various emotions in the Chinese culture. In English, the conceptualisation of emotions is also associated with a closed container, e.g. *pent-up anger/frustration/affection/feelings*. However, what makes the difference is that the Chinese concepts are specifically associated with the image of an organ as a container. Further, the Chinese culture associates them with specific sad feelings *yu* and *men*, and encodes these associations in its emotion lexicon; however, the English container conceptualisation is quite general as any feeling, whether positive or negative, can be *pent-up*.

Thirdly, Chinese and English differ in linguistic manifestations or specific conceptualisations of sadness, though they are similar in certain conceptual mappings between domains. For instance, *tongku*, *agony* and *anguish* are all conceptualised in the image of 'be in/into something' as we can find 'chu yu tongku zhi zhong' in Chinese, and 'in agony'/'in anguish' in English. However, Chinese *tongku* is also understood in terms of the 'falling into (*xianru*) something' image. These two images are quite close, and both are mapped from a container. The main difference between them is to do with

the ‘state’ character of the former and the ‘activity’ character of the latter. Various kinds of emotional states may be conceptualised as containers (Lakoff & Johnson, 1980; Machakanja, 2008) in both cultures. We can find ‘in love’, ‘in grief’, ‘in depression’, ‘in excitement’ in English and their counterparts ‘chuyu lian’ai/beitong/yiyu/xingfen zhizhong’ in Chinese. It is also common to say ‘fall into depression’ or ‘fall in love’. But can we infer metaphors in a systematic way across cultures? We will still feel weird if we hear ‘fall into agony or anguish’. Another example to illustrate this would be *upset*. Both cultures map the spatial orientation onto the feeling of sadness, but the association of the image of ‘overturn’ with sadness is English-specific. Considering its metaphorical process, it is not hard to understand that *upset* is a complex concept characterised by a mixture of *sadness*, *annoyance*, *worry* and *anxiety*. There is no such a Chinese word which incorporates so many emotional components into one simple concept just because there is no such a metaphorical construal of sadness in Chinese. The issue relevant to metaphorical words is that the difference in word formation will lead to their difference in usage patterns, so many so-called translation equivalents are not truly correspondent to each other. For example, the association of the image of *chongman* (being filled with) with an intense feeling seems to be shared by English and Chinese, maybe also other languages, but Chinese never uses *chongman* to describe *gaoxing* (happiness), *shangxin* or *nanguo*. They cannot be metaphorically understood in terms of ‘a full container’ probably because they are already metaphors which are associated with other images, so they cannot be used in this way.

Fourthly, even if sadness is related to the mapping between the same source domain and the same target domain in the two cultures, when it comes to the exact meaning it refers to, they exhibit some cultural differences. For example, although both cultures map the physical experience of injury onto the sad feeling, *shangxin* and *xinsui* in Chinese are both quite different from the English *heartbroken** in use and meaning. This can also be illustrated by the conceptualisation of *beitong*. Apart from *chongman*, we also use *chenjin* (be immersed into) to describe *beitong*. Although *chenjin* also presents us an image of ‘being full’, the association of a strong emotion with the state of ‘be immersed in’ is rather special to Chinese. So it can be argued that in distinctive cultures the specific associations arising from the same or similar conceptual mapping might be different and consequently lead to different linguistic manifestations. This seems to suggest that languages may display cultural variations and distinctiveness at varying levels of metaphorical conceptualisations. Therefore, cross-linguistic studies on metaphors should go deeper to look at their differences in specific conceptualisations, rather than stop at the level of mapping between domains.

Studies on emotions show that the metaphorical conceptualisations of emotions are closely related to the conceptualisations of body parts, which display considerable cultural variation (Ameka, 2002; Enfield & Wierzbicka, 2002; Turpin, 2002; Ye, 2002; Yu, 2002, 2003, 2009), though human beings undergo similar bodily events or processes in our emotional experience. One typical example would be that in the Chinese culture the organ *dan* ‘gallbladder’ is related to courage because the gallbladder is thought to play a role in making judgments and decisions and therefore has to do with courage (Yu,

2003). Chinese has traditionally linked various emotions with internal organs, which can be traced back to traditional Chinese medicine. An overview is given below, with the metaphorical images evoked in the expressions given in the parentheses (Yu, 2009: 239).

- (i) spleen: anger (gaseous energy of qi expanding)
- (ii) liver: anger (fire burning)
sadness (split, ripped into pieces)
- (iii) intestines: anxiety (knotted, twisted, hanged)
sadness (knotted, twisted, split, ripped into pieces)
- (iv) stomach: anxiety (hanged)
- (v) lungs: sadness (split)
- (vi) gallbladder: fear (void, frigid, hanged, dropped, lost, trembling, split)
- (vii) heart: anger (fire burning)
anxiety (troubled, hanged, lifted, suspended, pinched, scorched, fried, burned by fire)
sadness (pain, frigid, wounded, twisted, fragmented, pierced, torn apart, pounded)
fear (shocked, throbbing, frigid, lifted, split)
happiness (open, vast, bright, blooming)
relief (laid down, settled down, calm down, broad)
admiration (toppled, drunk)
vexation (closed up, blocked, messed up, troubled, pain)
disappointment (frigid, gray)
disgust (vomiting)
hatred (pain, rotten)
guilt (void, lost)

This list is based on dictionaries. These metaphorical images can roughly reflect how emotions are understood in the Chinese culture. We can see that the heart is especially rich in emotional associations, which explains why Yu (2009) makes a book-length analysis of *xin/heart* metaphors, in which he points out that ‘the dichotomies between body and mind, heart and head, in English do not exist in traditional Chinese thought’ (p. 352) and the heart has been traditionally seen as the organ of thinking and feeling in the Chinese culture. All these associations have been encoded in the Chinese lexicon. 131 words starting with *xin* are found in the XHCBL and the great majority of them are

metaphors, like *xinsheng* (heart-voice/heartfelt wishes), *xinbing* (heart-disease/worry), *xinyuan* (heart-wish/wish) and *xinxi* (heart-thin/cautious; careful). There are still numerous metaphorical *xin*-words not starting with *xin* which cannot be easily counted due to the searching method, such as *fangxin* (lay down-heart/rest assured), *jiaoxin* (burnt-heart/anxious), *yixin* (suspicious-heart/suspicion), *guanxin* (close-heart/concern;care), *naixin* (endure-heart/patience), *xinxin* (believe-heart/confidence), *juexin* (determine-heart/determination) and *xiongxin* (masculine-heart/ambition). The illustration of the *xin*-words might help understand why metaphors are pervasive in the Chinese lexicon as well as in the sadness lexicon (expressions).

Yu's analysis of *heart* is comprehensive and almost exhaustive, but as he states in the conclusion of the book, he mainly leans on a qualitative approach, which will inevitably lead to some unsatisfactory findings. Consider the following quotes about *shangxin* from his book (p. 259) :

- 1) 心碎 *xin-sui* (heart-fragmented) 'be heart-broken'
- 2) 伤心 *shang-xin* (wound-heart) 'sad; grieved; broken-hearted'

From the above quotes, it is clear that he thinks both *shangxin* and *xinsui* are the same as *heartbroken* or *broken-hearted*. But my corpus-based analysis in Section 5.2.4 shows that neither *shangxin* nor *xinsui* is the same as *heartbroken**: *shangxin* as a word in modern discourse denotes a more outward but less intense emotion than *heartbroken*, and *xinsui* is a more intense feeling than *heartbroken* but sometimes it could be absolutely positive. This divergence seems to suggest that the subtle differences of conceptual metaphors between distinctive cultures can only be revealed by corpus-based analyses. Another thing that should be noted in his qualitative analysis is that he

analyses many words that are hardly used in real language. For example, he discusses 心伤 and also glosses it as ‘*xin-shang* (heart-wounded) ‘sad; grieved; broken-hearted’ (p. 259). However, only 115 hits of *xinshang* are found in the CCL corpus and few of them are used as a word, as exemplified in the following lines:

她说: ‘我们的 心伤 透了’, 我们之所以加入国大党的阵营
战争对 130 多万名伊拉克儿童所造成的身 心伤 害远不止

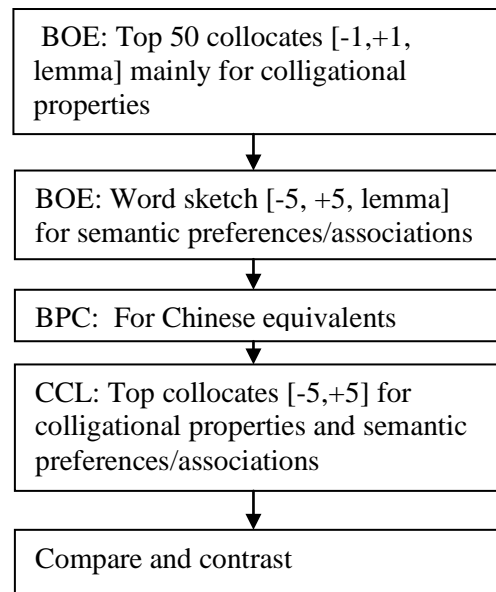
It is clearly seen that in ‘心伤透了’ (heart-pierced) and ‘身心伤害’ (physical and mental hurt) in the above lines, 心 and 伤 happen to co-occur, but they do not serve as a word. A qualitative analysis like this attaches equal weight to the words which occurs more than 2800 times and only a few times in a corpus, while a corpus-based analysis puts particular stress on frequency and real language use.

5.4 A summary of the corpus-based contrastive analysis

This section is a brief conclusion that sums up the major findings of this research, based on the contrastive analysis in Section 5.2. The steps my analysis followed can be schematically presented in Figure 5.2 (p. 257). Basically I followed these procedures. For some words, I also searched for their collocates in the span of -5 to +5. The research yields an equivalence network of sadness expressions between English and Chinese, which is presented in Figure 5.3 (p. 258).

Looking at the figure, our first impression is that of chaos. This is an equivalence network principally based on a unidirectional parallel corpus, i.e. from English to Chinese. From it we can see how these sadness expressions match up between English and Chinese. By carefully comparing the Chinese synonyms

Figure 5.2 Diagram for contrastive corpus-based analysis



obtained from the parallel corpus, we have figured out the distinction dimensions so we can distinguish them from each other. At the same time, the semantic information of the original English lexical item has been revealed. My intuition has also served to complement the research findings. For example, I added *yiyu*, a quite modern word, to the equivalent group of *depression** to compensate for the non-contemporary texts included in the parallel corpus.

It is true that there are some common points in Chinese and English sadness expressions. For one thing, they are similar in certain metaphorical mappings through which sadness is conceptualised, as noted earlier. For another, sadness expressions in both languages have strong colligations with modifiers and the verb *feel* (*ganjue/gandao/juede/jue*); both are semantically associated with words related to its physical manifestations, like *eyes, voice, facial expression, appearance*. Nevertheless, their cultural/linguistic differences observed in attested data seem more prominent than their commonalities. Table 5.13.2 (p.259-260) summarises the major findings concerning the semantic differences of the examined lexical items:

Figure 5.3 Equivalence network of sadness expressions between English and Chinese

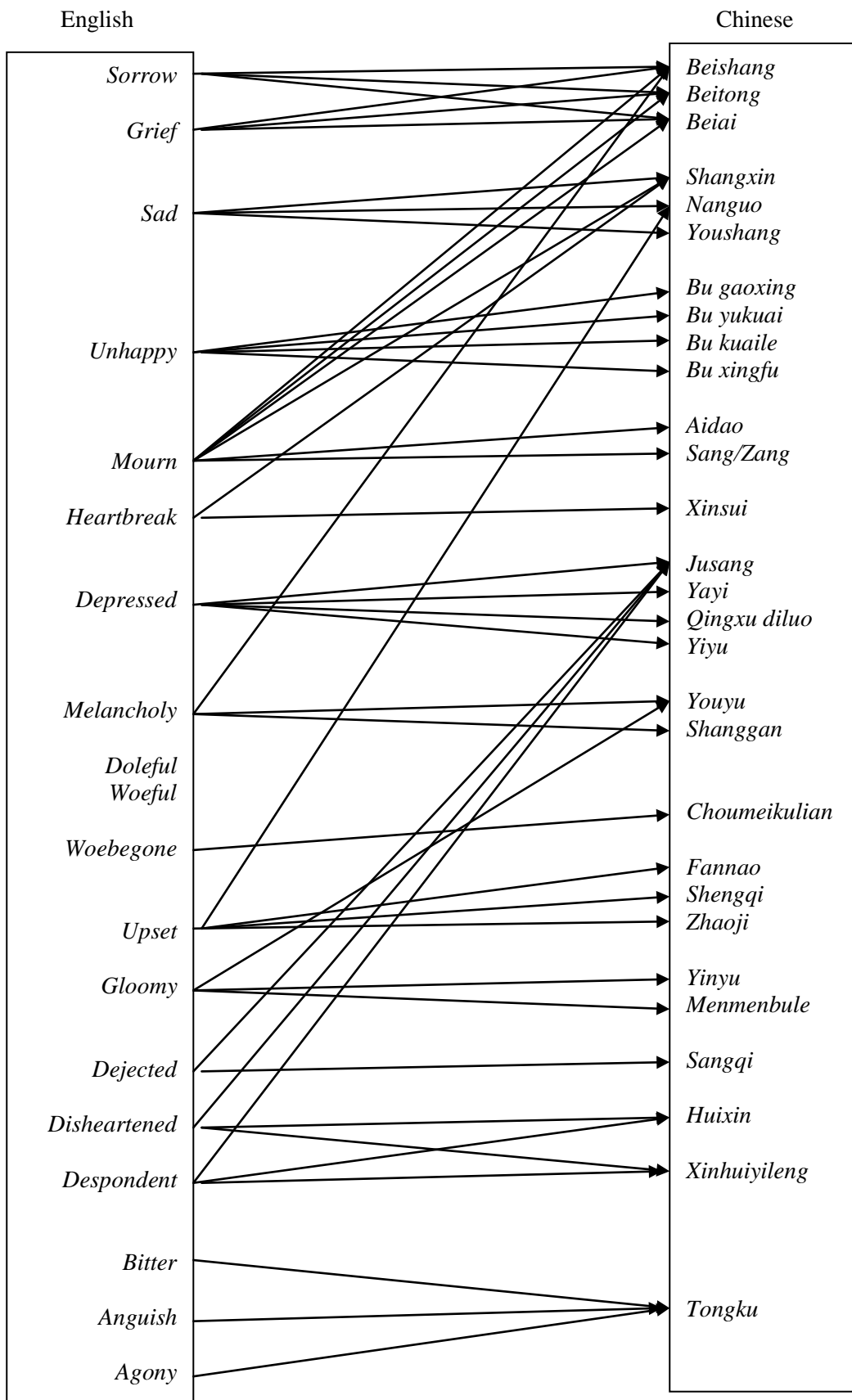


Table 5.13. 2 Major differences between English and Chinese sadness expressions

Words	English	Chinese
sorrow/grief <i>beishang/beiai/ beitong</i>	lacking the dimension: altruistic or personal	Altruistic—Chinese specific
unhappy <i>bu gaoxing/bu kuaile/bu yukuai/ bu xingfu</i>	<i>Unhappy</i> covers the meanings of all four Chinese equivalents	Dimensions: introversion/extroversion, specificity/generalality, length
sad <i>shangxin/ nanguo/ youshang</i>	<i>Sad</i> seems to be a generic term. It means ‘feeling unhappy’, ‘showing unhappy’ and ‘making someone unhappy’.	No real generic term in modern Chinese. Compared with <i>sad</i> , <i>shangxin</i> is more intense, specific and extroverted, <i>nanguo</i> is also more specific, and <i>youshang</i> is more intense, more literary, more introverted and less negative. <i>Nanguo</i> may be caused by oneself or someone else, while <i>shangxin</i> is more likely to emphasise the pain on ‘ME’ caused by other people or some external force.
heartbreak <i>shangxin/xinsui</i>	<i>Heartbreak</i> lies between <i>shangxin</i> and <i>xinsui</i> in terms of intensity.	<i>Shangxin</i> seems to be more extroverted and <i>xinsui</i> is mainly used in causative constructions and less negative.
mourn <i>beishang/beitong/ beiai/shangxin/ aidao</i>	<i>Mourn</i> : 1) feel sad and miss someone after his/her death, 2) show publicly one’s bereavement that may or may not be sincere, based on ritualised or social practice, and 3) feel regret about the loss of something	<i>Aidao</i> : show one’s sadness about someone’s death, based on ritualised or communal practice <i>beishang/beitong/beiai/shangxin</i> : voluntary feeling of sadness
doleful/woeful/ woebegone	Mainly showing one’s sadness	No exact equivalent. Chinese does not distinguish between feeling and showing sadness in the lexicon. It tends to combine emotion words and facial organs to describe emotional manifestations.
depressed/ melancholy <i>jusang/yayi/ qingxu diluo/yiyu</i> <i>beishang/youyu/ shanggan</i>	<i>Depressed</i> has a wider semantic range, covering the meanings of <i>jusang</i> , <i>yayi</i> , <i>qingxu diluo</i> and <i>yiyu</i> . <i>Melancholy</i> and <i>youyu</i> are similar in that they are both related to a temperament or character and sometimes can be used in a less negative way. But it is closer to <i>shanggan</i> in feeling sadness for no particular reason.	Each equivalent emphasises a different aspect. <i>Jusang</i> focuses on disappointment, frustration and hopelessness and <i>yayi</i> on the difficulty or impossibility of untying the knots and the incompetence to release one’s sad feelings. <i>Yiyu</i> is closer to <i>depression</i> , but not <i>depressed</i> . <i>Youyu</i> is much less negative. <i>Shanggan</i> is the sadness invoked by emotional stirrings.
upset <i>nanguo/fannao /shengqi/zhaoji</i>	<i>Upset</i> is complex concept characterised by a mixture of several feelings, i.e. <i>sadness</i> , <i>worry</i> , <i>anxiety</i> and <i>annoyance</i> .	No exact Chinese equivalent. Chinese does not encode so many components into one single emotional concept. It seems that Chinese emotional concepts tend to be more discrete, with each emphasising one emotional aspect.
gloomy <i>youyu/yinyu/ menmenbule</i>	<i>Gloomy</i> is more semantically associated with the future. The past vs future dimension is lacking in Chinese <i>gloomy</i> words.	<i>Youyu</i> is more an inward feeling without specific reasons. However, sometimes it is also viewed as something very beautiful. <i>Yinyu</i> is more manifested outwardly. Both <i>youyu</i> and <i>yinyu</i> can be related to a person’s character. People will feel <i>menmenbule</i> when things are not going as wished.

dejected/ disheartened/ despondent <i>jusang/sangqi/ huixin/ xinhuiyileng</i>	<i>Dejected</i> focuses on the reaction to what has just happened, while <i>despondent</i> emphasises the feeling due to the gloomy prediction about the future, and <i>disheartened</i> centres on both.	<i>Sangqi</i> seems to focus on the emotional reaction to what has just happened, but no Chinese word emphasises the pessimistic thinking about the future; <i>huixin</i> and <i>jusang</i> focus on both, with the former more introverted and the latter more extroverted.
bitterness/ agony/ anguish <i>tongku</i>	<i>Bitterness</i> indicates an inner emotional state, and it is equivalent to <i>tongku</i> only in the non-opposition sense. <i>Anguish</i> implies extreme pain or suffering and is hardly modified by downtoners. <i>Anguish</i> is more conceptualised as a static emotional state, while <i>agony</i> is more understood as a dynamic process. Of these words, <i>agony</i> has the broadest range of semantic coverage, followed by <i>anguish</i> and <i>bitterness</i> .	<i>Tongku</i> has a narrower semantic coverage than <i>agony</i> and is less intense <i>anguish</i> . It talks about the suffering invoked by something specific, especially if one is not able to get out of it due to the incompetence to untie the mental knot.

Only the relatively important findings are presented in the above table because the semantic details are too rich to be listed here. I will sum up the broader linguistic/cultural differences revealed by the earlier contrastive analyses as follows:

- 1) In terms of colligational pattern, firstly, Chinese sadness expressions, different from their English counterparts, have a strong colligation with prepositional phrases which indicate the perceived loci related to the inner emotional or outer physical activities, typically *xinzhong/xinli* (in the heart) and *lianshang/manlian* (on the face). This might be because Chinese does not distinguish between feeling and showing sadness in the lexicon. Instead, it uses the collocation with other words, typically prepositional phrases, such as ‘in the heart’, ‘on the face’, to express the difference. Secondly, they tend to be preceded by classifiers, such as *zhong* (kind), *fen* (portion), *si* (lit. silk/minute), *ci* (times) and *zhen* (lasting for a short time). The Chinese classifier system is not only a system of linguistic categorisation, but also reflects its underlying conceptual structure’ (Jiang, 2004; Tai, 1994; Tai & Chao, 1990). For example, *si* is an extension from the shape of long, thin silk to a slight intensity of feeling. Thirdly, unlike their English counterparts, Chinese sadness expressions strongly colligate with the particle *de*, but hardly with various forms of *be*, which, of course, is due to the typological differences between these two languages.

- 2) In terms of collocation, firstly, Chinese sadness expressions are more semantically associated with *biaoqing/shense/shenqing* (facial expression) and *yangzi* (appearance), which are realised by link verbs in English, mainly *look*, *appear*, *seem* etc. Secondly, although sadness

expressions often collocate with their hypernyms in both Chinese and English, they collocate with different superordinates. In English, most sadness expressions, such as *sadness* (148, -5, +5), *grief* (83, -5, +5), *sorrow* (21, -5, +5), often co-occur with *feeling*; in contrast, in Chinese, most sadness expressions frequently co-occur with *qingxu* (mood) or *xinqing* (frame of mind), rather than the Chinese translation equivalents of *feeling--qinggan* or *ganqing*. This is probably because *qinggan* or *ganqing* is not truly equivalent to *feeling* (R. Zhang & Ooi, 2008). It seems that in the Chinese discourse *sadness*, as well as other feelings, is more conceptualised or talked about as *qingxu* or *xinqing*, rather than *ganqing* or *qinggan*. Thirdly, Chinese sadness expressions collocate more frequently with their hypernyms—*qingxu* or *xinqing* ---than their English counterparts do with *feeling*.

- 3) Unlike English, Chinese sadness expressions are predominantly preceded by causative constructions, such as *ling/shi/rang (ren)*, to express the meaning of ‘making or causing someone sad’, while such a meaning has been incorporated as a sense into the English sadness lexicon. This suggests that in expressing causative meanings, Chinese sadness expressions rely more on analytic expressions. Further, this feature seems to be applicable to the whole language (cf. Xu, 2002). So we can argue that, at least in terms of causative expressions, English has a higher degree of lexicalisation than Chinese.
- 4) For metaphorical sadness expressions, Chinese and English differ in certain conceptual mappings between domains. For example, as noted earlier, Chinese maps the domains of colour, temperature, weather, living a life onto the domain of sadness, which English lacks; however, unlike English, Chinese does not associate sadness with darkness.
- 5) Chinese and English differ in linguistic manifestations or specific conceptualisations of sadness, though they are similar in certain conceptual mappings between domains. A typical example for this is that both Chinese and English construe sadness as a container, but the English construal is rather general, and the Chinese understanding, by contrast, is based on a specific system relating various emotions to different organs. This is much to do with the traditional Chinese medicine.
- 6) Many modern Chinese sadness expressions are still related to the traditional fatalism as many Chinese people, especially at the time of difficulties, still believe that events are decided by fate and there is nothing one can do to prevent them from happening. If one feels that they cannot change or control the fate, s/he thinks it is pointless to try.

Beishang, *shanggan*, *youshang* and *youyu* are all relevant to the fatalistic view, to different extents.

- 7) *Sad* is an umbrella term for all sadness expressions in English, while modern Chinese lacks such a cover term. The closest one would be *shangxin* or *nanguo*, but neither of them fulfils the same function as *sad* does in English. We can also find that *unhappy* covers a broad semantic range and it incorporates into one single word the meanings of Chinese *bu xingfu*, *bukuaile*, *bu yukuai*, *bu gaoxing*, *buxing*. *Unhappy* has a wide semantic range because *happy* is also a cover term like *sad*. Modern Chinese lacks such a cover term probably because it is a mixture of ancient Chinese characters and derived modern words. Each ancient Chinese character has a distinctly clear meaning, such as *bei* and *xi*. In modern Chinese, they can be used as a cover term, but often in combinations, like *bei xi jiaojia* (sorrow+joy+mingle) or *xi nu ai le* (happiness+anger+sadness+joy). This might have hindered the modern cover terms in the emotion domain, probably in other domains as well, from coming into being in modern Chinese.
- 8) Chinese and English apply different dimensions to distinguish sadness expressions. For instance, in Chinese, the dimension of voluntary vs ritualised is applied in distinguishing Chinese sadness expressions, in which *shangxin/nanguo/beitong/beishang* is voluntary and *aidao* is ritualised and culturally mediated. There is a distinct division of labour between them. In English, there is no such a clear dividing line. For example, *mourn* combines both voluntary and ritualised feelings into one word, though *mourning* seems to have a strong tendency to be more ritualised, but it covers three Chinese concepts *aidao*, *sang* and *zang*, which are also distinctly distinguished. This point can also be illustrated by the concept of *beiai*, which is unique for its altruistic characteristic. The dimension of *personal* (TO ME) vs *altruistic* (TO OTHERS) is Chinese specific. Further, the Chinese *gloomy* group lacks the dimension of past vs future and so does the Chinese *dejected* group. Languages differ in ways to combine semantic components into words; in other words, different languages differentiate their words or synonyms along different dimensions, which reflect the ways in which their speakers organise, classify and categorise the world.
- 9) Some Chinese sadness expressions are often associated with a positive or less negative semantic prosody, as exemplified by *xinsui*, *youyu* and *youshang*. The case of *xinsui* shows that the same image can be understood in different ways in different contexts within the same language. The cases of *youyu* and *youshang* might be to do with the constituent character *you*, which was traditionally associated with

wisdom, knowledge, contemplation, far-sightedness etc. *You* still keeps its less negative flavour in many Chinese words.

- 10) It seems that English sadness expressions have a wider semantic space than their Chinese counterparts because quite a few English sadness expressions such as *depression*, *gloomy* are also associated with economy, but Chinese sadness expressions are only confined to emotions and seldom used to talk about economy or other things, except that they are used in their original senses.

I have discussed, in detail, the similarities and differences displayed in the English and Chinese sadness expressions. It seems that almost all of the above findings can be generalised to the whole emotion domain of these two languages. It is also possible for some of them, such as (3), (4), (5), (7), (8), to be extended to other domains. The research findings appear to suggest that the cultural variations are best explored and described using corpus methodology as the subtle nuances of meaning between semantically similar expressions can only be discerned by looking at their colligations, collocations, semantic associations as well as a large number of examples. The corpus data can provide not only hard evidence for the analyst's introspection but also new findings that are less accessible to human intuitions.

CHAPTER 6

IMPLICATIONS

6.1 Introduction

The previous chapter contrasted the colligations, collocations and semantic associations of near synonyms in the emotion domain between English and Chinese by drawing on corpus data from two large monolingual corpora and a parallel corpus. This chapter will address potential applications of this methodology in the areas of psychology, contrastive lexical studies, bilingual lexicography and language pedagogy.

6.2 Implications for related research areas

6.2.1 Psychology

As mentioned in the introduction of this thesis, the psychological approach to the study of emotions lacks objectivity for two reasons. For one thing, such an approach is merely based on introspection. For another, the subjects studied are unlikely to behave in a natural way under observation. The corpus-linguistic approach to the study of emotions, as illustrated in this research, does not have such problems. It is based on texts written independently of the study, so the results are less distorted by the subjects. Besides, accuracy and reliability can be achieved because the number of texts examined is much larger than could be looked at in psychological studies. So using corpora to study emotions has at least two advantages. Firstly, the study bases its conclusions on actual usage rather than on mere introspection. Secondly, because a corpus is a collection of numerous opinions about the question under investigation, the corpus-based study can

consider a large number of opinions without interviewing informants. The corpus-based contrastive research on emotions can reveal the multiple perspectives on human emotional experience reflected in ‘emotion’ terms in different languages and show human emotions are neither universal nor innate, which has been mistakenly claimed by many psychologists. In fact, the corpus-based approach can also be extended beyond emotions to other domains in psychology.

6.2.2 Contrastive lexical studies

Previous studies of cross-linguistic lexical semantics have been usually conducted in the field of traditional contrastive lexical semantics, cognitive linguistics, cultural linguistics or Natural Semantic Metalanguage. In this research I have developed a unique theoretical approach to compare and contrast lexical meaning, which integrates corpus-linguistic theories on meaning (as a social construct, usage and paraphrase) (Teubert, 2005, 2010; Teubert & Cermakova, 2004) and corpus-linguistic categories, i.e. colligation, collocation and semantic preference/association (Sinclair, 1996; Hoey, 2005). A new complex methodology has also been developed to tease meaning out of corpus evidence.

It is a new attempt to extend the corpus-linguistic framework to the field of contrastive lexical studies. The research results suggest that this approach, compared with other approaches basing analyses mainly on inspection, qualitative reasoning and intuition, can present more details about word usage patterns and meaning. It seems to be a more viable way to distinguish the subtle nuances of meaning between synonyms within and across languages. The use of parallel corpora to establish the correspondence

between linked expressions can offer quantitative information on the congruence between two languages. One reason for the revival of contrastive linguistics is the possibility of using large corpora (Altenberg & Granger, 2002). The strength of using corpora in contrastive lexical semantics should not be underestimated. In addition, it is also possible to use the corpus-based methodology to contrast languages at other levels, such as phraseology, syntax, grammar.

6.2.3 Bilingual lexicography

The methodology of the present work can shed some new light on bilingual lexicography. Most old dictionaries were based on ‘a mixture of citations, introspection, and what other dictionaries said’ (Moon, 2007: 7). Monolingual lexicography, since the first publishing of Sinclair’s *Collins Cobuild English Language Dictionary* (CCELD) in 1987, has entered a new era, in which a dictionary has become ‘a record of how language was actually used, not thought to be used, with this record constructed from corpus evidence’ (Moon, 2007: 7). Its birth had a massive impact on the practice of commercial dictionary-making, lexicographical theory and language description in general (Moon, 2007). The advent of corpora not only provides a more objective source of evidence, but also leads to a new look at language (Moon).

Bilingual lexicography has far lagged behind monolingual lexicography in terms of the application of corpora. Take English-Chinese for example. After 24 years, no English-Chinese dictionaries comparable with COBUILD dictionaries have come out. The Longman and Oxford English-Chinese dictionaries have been revised several times, but they are still based on the old lexicographical theory and principles. As shown in

earlier sections, the Chinese definitions in those dictionaries are simply translations of their English parts. Therefore, there is a pressing need for the compilation of corpus-based bilingual, especially English-Chinese, dictionaries, in which the detailed information about synonyms in English should be included. Such information can make language learners aware of the distinctions between synonyms within and across languages. Of course, the implementation of such a project will pose some problems. One is that it is hugely time-consuming to analyse corpus data. Another is how to organise the analysis results, i.e. how to present the colligational and collocational patterns and the information on the semantic associations of words within limited space. An electronic dictionary might be a viable option.

As translational equivalents between L1 and L2 do not match up neatly, the traditional practice of glossing L2 words by offering single L1 words should be used with caution. Language learners wish to know more details about how to use a particular word in specific situations. It is impossible for learners to learn every word by reading concordances, albeit possible for some words. The details concerning repeated patterns, actual uses and exact meanings should be made available in the bilingual dictionary. It does not matter in whatever form to present the information on word use and meaning. What does matter is that language learners can find such information somewhere. It will save learners a lot of effort and time.

6.2.4 Language pedagogy

This section addresses the potential uses of the corpus-based methodology in the English language classroom. It has two implications for language pedagogy, both related to

vocabulary teaching, particularly the teaching of synonyms in L2. For one thing, translational equivalents across languages were assumed to be unproblematically 'the same', but the corpus-based analysis of this research reveals that apparently similar items in fact display appreciable differences. Therefore, the traditional practice of using translational equivalents in L1 to explain L2 vocabulary should be used with caution. Relating one single word in L1 to another in L2 will mislead learners into believing they are the same. Language learners need to be made aware of the disparity between what a word exactly means and what can be found in its gloss in the textbook.

For another, inappropriate word choice arising from a partial understanding of its meaning is quite common in second/foreign language learning. In-depth knowledge about the distinctions between semantically similar words would be highly useful to L2 learners. Native speakers can detect the usage of words and use them correctly, based on their intuition, albeit they are less able to describe them. However, language learners have to rely much on the information from teachers or dictionaries to fully grasp them, especially synonyms. It is quite common that even teachers cannot tell how to distinguish them, although they have no difficulty in using them. If teachers can use corpora to analyse and contrast synonyms in L2 and their translational equivalents in L1, summarise their findings about to what extent they are similar and how they differ, and show them to learners, their production errors will be greatly reduced.

CHAPTER 7

CONCLUSIONS

7.1 Summing up

The concluding chapter will briefly sum up the whole research project. Chapter 1 offered an introduction to the research, in which the impetus, methodology, significance and aims of this research were given. Chapter 2 reviewed relevant literature on emotions in psychology and linguistics, contrastive lexical studies and corpus linguistics. Chapter 3 addressed the theoretical framework of the research and Chapter 4 dealt with its methodology. The corpus research findings yielded were presented in Chapter 5, along with a complementary contrastive-qualitative analysis of metaphorical sadness expressions in English and Chinese. Chapter 6 discussed the implications of this research for research areas of psychology, contrastive lexical studies, bilingual lexicography and language pedagogy. This chapter will sum up the present work, discuss the limitations of the study and finally propose some further research topics.

This research deals with a cross-linguistic contrast of sadness expressions between English and Chinese. An in-depth analysis has been made on the basis of the corpus data by applying corpus-linguistic theories on meaning, i.e. meaning is a social construct, usage and paraphrase, and by employing a corpus-linguistic lexical model, i.e. to describe lexical items in terms of colligations, collocations and semantic preferences/associations. The approach used is embedded in a corpus theoretical framework. It incorporates the use of parallel corpora into this methodology to establish the lexical links across languages and provide illustrative examples.

This research extends the corpus-linguistic framework to the field of contrastive lexical semantics between English and Chinese. It gives fresh insight into how developments in corpus-linguistic theories and in the application of computer corpus technology can advance the field of lexical contrastive linguistics. It is a nuanced description of similarities and differences that epitomises the descriptive potential of the corpus-based approach to contrastive lexical semantics.

Apart from a quantitative analysis of sadness expressions based on the corpus data, a qualitative analysis of metaphorical sadness expressions in English and Chinese has also been made. The analysis reveals some similarities in the way in which these two cultures conceptualise emotions metaphorically, as well as some remarkable differences in specific cultural conceptualisations and linguistic manifestations. The detailed corpus-based analysis shows that the cultural disparities reflected at the lexical level of language also come from two other sources. One is from the different ways in which two languages encode meanings in their lexicons and the different dimensions used to distinguish words/synonyms in distinctive cultures. The other would be the influence from their traditional philosophies when meanings were encoded in the lexicon at the outset. After the development for a long time, there are still discernible traces in their meanings.

The research findings clearly show that human emotions are not innate nor universal, but culture-specific, which is contrary to previous psychological findings. It seems that this corpus-linguistic approach can open up new possibilities for psychological studies on emotions. Contrastive lexical studies can also benefit from this

approach in that the subtle nuances of meaning can be captured and consequently semantically similar items within and across languages can be distinguished. Not only can what they share and how they differ be discovered, but the impact of cultural differences on encoding meaning in the lexicon can be revealed. By analysing corpus data, cultural conceptualisations, whether metaphorical or literal, can be highlighted and the understanding of how emotions are understood in distinctive cultures can be greatly enriched. In addition, the application of this approach to bilingual lexicography and language pedagogy should not be underestimated.

7.2 Limitations

The methodology of corpus linguistics uses directly observable language data and computational tools for the detection and classification of linguistic phenomena. Both data and computer programs are open to scrutiny by the peer community. The results they produce are reproducible. However, we also have to be aware that corpus linguistics has its limitations. Firstly, a corpus-based analysis is time-consuming, and consequently the number of linguistic items that can be handled in a study is limited. Secondly, it is true that corpus data are objective, but the interpretation of the data may not be totally so because when the analyser interprets the data, s/he will unavoidably add a subjective point of view to the analysis. Even in corpus tagging or marking, human beings are reproducing their bias (Teubert, in Zhang, 2009). This is why Sinclair (1991) was opposed to corpus annotation.

This research is concerned with cross-linguistic contrast of sadness expressions in English and Chinese. For this research, there are several limitations. First, the

monolingual corpora I used were not exactly comparable in size and composition. Second, the BPC used to establish the cross-linguistic congruence was made up of only fictional texts, so the Chinese equivalents retrieved from this parallel corpus might not be exactly the same as those from a general corpus. Third, the size of the BPC is rather limited, containing only a total of 3,707,397 tokens. Similarly, the Chinese equivalents found based on this small parallel corpus might differ, maybe slightly, from those using a large parallel corpus. It is for the readers to decide to what extent the research findings would differ if a large general English-Chinese parallel corpus had been used. Fourth, the Chinese data were not POS tagged. Although it is still arguable whether POS tagging is encouraged, whether Chinese has the same POS classification system as English, if we want to compare and contrast English and Chinese quantitatively, it seems to be better to have both POS tagged on the basis of function so that they can be comparable. Fifth, the English collocates were produced by Sketch Engine based on t-score, while the Chinese collocates were obtained on raw frequency because the whole CCL corpus was not available and I could only get the concordance lines I needed. Without the corpus, it was impossible to produce collocates according to t-score.

I hope that future contrastive lexical studies can be done on the basis of a large general English-Chinese²¹ parallel corpus containing bilingual texts both produced after 1919 and two large monolingual corpora comparable in size²² and composition²³. It will

²¹ An English-Chinese/Chinese-English parallel (bi-directional) corpus would be more desired, but I am fully aware that it is less likely to be achieved because the translated texts from Chinese into English are much more limited than those from English into Chinese.

²² Comparability is only a matter of extent. Before we solve the problem of the correspondence between English words and Chinese characters, we cannot assure the total comparability in size.

²³ The ideal composition of a generous corpus is still controversial.

be better if any Chinese POS tagging software based on a well-defined tagset becomes publically available or any freely online available large Chinese corpus can generate data tagged by such a program. More reliable research findings will be achieved if the online concordancer of a large Chinese corpus can produce collocates based on t-score or even offer the Word Sketch and Sketch Diff functions like Sketch Engine. If all of these matters are resolved, we will get closer to the linguistic truth in the field of cross-linguistic lexical studies between English and Chinese.

7.3 Further research

The thesis has explored a new path in conducting contrastive lexical studies by adopting a corpus-linguistic approach which combines corpus-linguistic structural categories with corpus-linguistic theories on meaning. The research findings suggest a number of starting points for further research. Firstly, more in-depth studies on other emotions can be explored to discover more about cultural discrepancies reflected in the lexicon and to further support the claims made in this research. Secondly, studies on non-emotion words can be conducted to explore deeply about what the Chinese and English lexicons share and how they differ. As discussed earlier, for a particular language pair, the metaphorical mappings might be quite similar, but the linguistic manifestations and the exact meanings they refer to might not be so. Only after extensive research on words in more domains has been done, can claims about specific differences be made and consequently can a more complete picture concerning the Chinese and English lexicons be drawn. The dimensions and metaphorical mappings used in this study may not apply to other domains, but they can serve as a reference for future researchers to find new

dimensions and new mappings appropriate to their investigations. Thirdly, future focus can be extended from single words to bigger units, such as lexical items or (extended) units of meaning, so that Sinclair's (1996) lexical model, i.e. to investigate units of meaning in terms of colligation, collocation, semantic preference and semantic prosody, can serve as a viable tool to undertake this task. Fourthly, for other language pairs, it would be possible to proceed in a similar way. Fifthly, this corpus-linguistic approach to contrastive lexical studies can be further extended to language pedagogy as synonyms are of great importance in foreign/second language learning and teaching. A cross-linguistic comparison from the perspective of language pedagogy, with more emphasis on collocational behaviour and syntactic properties of words, might be highly helpful for language learners.

BIBLIOGRAPHY

A. Dictionaries

- Collins COBUILD Advanced Learner's English Dictionary*. (2006). (5th ed.). London/Glasgow: HarperCollins.
- Hanyu Da Zidian*. (1993). Wuhan/Chengdu: Hubei Cishu Chubanshe/Sichuan Cishu Chubanshe.
- Longman Dictionary of Contemporary English*. (2006). (4th ed.). Harlow, Essex: Pearson/Longman.
- Oxford Advanced Learner's English-Chinese Dictionary*. (2002). (4th ed.). Beijing: Shangwu Yinshu Guan.
- Oxford English Dictionary Online*. (2011). Oxford: Oxford University Press. Available at <http://www.oed.com.libproxy1.nus.edu.sg/>.
- Xiandai Hanyu Cidian (A Dictionary of Modern Chinese)*. (2005). (5th ed.). Beijing: Shangwu Yinshu Guan.
- Xiandai Hanyu Cidian: Bilingual Edition (A Dictionary of Modern Chinese)*. (2002). Beijing: Waiyu Jiaoxue Yu Yanjiu Chubanshe.

B. Other literature

- Aarts, J. (1998). Introduction In S. Johansson & S. Oksefjell (Eds.), *Corpora and cross-linguistic research* (pp. ix-xiv). Amsterdam: Rodopi.
- Aijmer, K. (1998). Epistemic predicates in contrast. In S. Johansson & S. Oksefjell (Eds.), *Corpora and cross-linguistic research* (pp. 277-296). Amsterdam: Rodopi.
- Aijmer, K., & Altenberg, B. (1996). Introduction. In K. Aijmer, B. Altenberg & M. Johansson (Eds.), *Languages in Contrast: Papers from a symposium on text-based cross-linguistic studies* (pp. 11-16). Lund: Lund University Press.
- Altenberg, B. (1999). Adverbial connectors in English and Swedish: Semantic and lexical correspondence. In H. Hasselgard & S. Oksefjell (Eds.), *Out of corpora: Studies in honour of Stig Johansson* (pp. 249-268). Amsterdam; Atlanta: Rodopi.
- Altenberg, B., & Granger, S. (2002). Recent trends in cross-linguistic lexical studies. In B. Altenberg & S. Granger (Eds.), *Lexis in Contrast* (pp. 3-50). Amsterdam: Benjamins.
- Ameka, F. K. (2002). Cultural scripting of body parts for emotions: On 'jealousy' and related emotions in Ewe. *Pragmatics and Cognition*, 10(1/2), 27-55.
- Aston, G. (1999). Corpus use and learning to translate. *Textus*, 12, 289-314.
- Averill, J. R. (1980). A constructivist view of emotion. In R. Plutchik & H. Kellerman (Eds.), *Emotion: Theory, research and experience* (pp. 305-339). New York: Academic Press.
- Baker, M. (1993). Corpus linguistics and translation studies: Implications and applications. In M. Baker, G. Francis & E. Tognini-Bonelli (Eds.), *Text and technology: in honour of John Sinclair* (pp. 233-252). Amsterdam: Benjamins.
- Baker, M. (1995). Corpora in translation studies: An overview and some suggestions for future research. *Target*, 7(2), 223-243.

- Bamberg, M. (1997). Culture, words and understanding. *Culture and Psychology*, 3(2), 183-194.
- Banczerowski, J. (1980). Some contrastive considerations about semantics in the communication process. In J. Fisiak (Ed.), *Theoretical issues in contrastive linguistics*. Amsterdam: Benjamins.
- Barlow, M. (2008). Parallel texts and corpus-based contrastive analysis. In M. Gómez González, L. Mackenzie & A. González (Eds.), *Current trends in contrastive linguistics: functional and cognitive perspectives* (pp. 101-121). Amsterdam: Benjamins.
- Bednarek, M. (2008). Semantic preference and semantic prosody re-examined. *Corpus Linguistics and Linguistic Theory*, 4(2), 119-139.
- Berger, & Luckmann, T. (1966). *The social construction of reality: A treatise in the sociology of knowledge*. Garden City, NY: Doubleday.
- Bhat, D. N. S. (1994). *The adjectival category: criterion for differentiation and identification*. Philadelphia: Benjamins.
- Biber, D., Conrad, S., & Reppen, R. (1998). *Corpus linguistics*. Cambridge: Cambridge University Press.
- Boucher, J. D., & Carlson, G. E. (1980). Recognition of facial expression in three cultures. *Journal of cross-cultural psychology*, 11(3), 263-280.
- Bowerman, M. (1989). Learning a semantic system: What role do cognitive predispositions play? In M. L. Rice & R. L. Schiefelbusch (Eds.), *The teachability of language* (pp. 133-169). Baltimore: Brooks.
- Bowerman, M. (1996). The origins of children's spatial semantic categories: Cognitive vs. linguistic determinants. In J. J. Gumperz & S. C. Levinson (Eds.), *Rethinking linguistic relativity* (pp. 145-176). Cambridge: Cambridge University Press.
- Brants, T. (2000). *TnT-- a statistical part-of-speech tagger*. Paper presented at the 6th Applied Natural Language Processing Conference. Retrieved from <http://www.coli.uni-saarland.de/~thorsten/publications/Brants-ANLP00.pdf>
- Briggs, J. L. (1970). *Never in anger: Portrait of an Eskimo family*. Cambridge: Cambridge University Press.
- Burr, V. (1995). *An introduction to social constructionism*. London: Routledge.
- Casad, E., & Langacker, R. (1985). 'Inside' and 'outside' in Cora grammar. *International Journal of American Linguistics*, 51(3), 247-281.
- Chang, L., Chen, K., & Huang, C. (2000). Alternations across semantic fields: A study on Mandarin verbs of emotion. *Computational linguistics and Chinese Language Processing*, 5(1), 61-80.
- Chao, Y. R. (1968). *A Grammar of Spoken Chinese*. Berkeley: University of California Press.
- Chesterman, A. (1998). *Contrastive functional analysis*. Amsterdam: Benjamins.
- Cheung, M. L. L. (2009). *Collaborative knowledge construction*. PhD Thesis. University of Birmingham.
- Chief, L., Huang, C., Chen, K., Tsai, M., & Chang, L. (2000). What can near synonyms tell us? *Computational linguistics and Chinese Language Processing*, 5(1), 47-60.
- Choi, S., & Bowerman, M. (1992). Learning to express motion events in English and Korean: The influence of language specific lexicalization patterns. In B. L. S. Pinker (Ed.), *Lexical & conceptual semantics* (pp. 83-121). Cambridge; Oxford: Blackwell.
- Chomsky, N. (1986). *Knowledge of language*. New York: Praeger.

- Darwin, C. (1872/1965). *The expression of the emotions in man and animals /with a preface by Konrad Lorenz*. Chicago: University of Chicago Press.
- Dixon, R. M. W., & Aikhenvald, A. Y. (Eds.). (2004). *Adjective classes: A cross-linguistic typology*. Oxford; New York: Oxford University Press.
- Durkheim, E. (1911/1961). *The elementary forms of the religious life*. New York: Collier Books.
- Dyvik, H. (1998). A translational basis for semantics. In S. Johansson & S. Oksefjell (Eds.), *Corpora and cross-linguistic research* (pp. 51-86). Amsterdam: Rodopi.
- Ekman: (1973). *Darwin and facial expression: A century of research in review*. New York: Academic.
- Ekman: (1993). Facial expression and emotion. *American Psychologist*(48), 384-392.
- Ekman: (1994). Strong evidence for universals in facial expressions: A reply to Russell's mistaken critique. *Psychological Bulletin* 115(1), 102-141.
- Ekman: (2007). *Emotions revealed: Recognizing faces and feelings to improve communication and emotional life [1st ed. 2004] (2nd ed.)*. New York: Henry Holt.
- Ekman., & Davidson, R. J. (1994). *The nature of emotion: fundamental questions*. Oxford; New York: Oxford University Press.
- Ekman., & Friesen, W. V. (1975). *Unmasking the Face: A guide to recognizing emotions from facial clues*. New Jersey: Prentice Hall. (Reprinted, Palo Alto, Calif.: Consulting Psychologists Press, 1984.).
- Ekman., Sorensan, E. R., & Friesan, W. V. (1969). Pan-cultural elements in facial display of emotions. *Science*, 164, 86-88.
- Elfenbein, H. A., & Ambady, N. (2002). On universality and cultural specificity of emotion recognition. *Psychological Bulletin*, 128(2), 205-235.
- Enfield, N. J., & Wierzbicka, A. (2002). Introduction: The body in description of emotion. *Pragmatics and Cognition*, 10(1/2), 1-25.
- Fairclough, N. (1989). *language and power*. London: Longman.
- Fillmore, C. J. (2003). *Form and meaning in language: papers on semantic roles*. Leland, CA: CSLI Publications, Center for the Study of Language and Information, Leland Stanford Junior University.
- Fillmore, C. J., Johnson, C. R., & Petruck, M. R. L. (2003). Background to FrameNet. *International Journal of Lexicography*, 16(3), 235-250.
- Firth, J. R. (1957). *Papers in linguistics* (pp. 168-205). Oxford: Oxford University Press.
- Francis, G., Hunston, S., & Manning, E. (1996). *Grammar patterns 1: Verbs*. London: HarperCollins.
- Gao, H. (2009). Reexamining multiple class membership in Chinese from a cognitive grammar perspective. *Chinese Language Learning*(2), 17-24.
- Garfinkle, H. (1967). *Studies in ethnomethodology*. Englewood Cliffs: Prentice Hall.
- Gellerstam, M. (1986). Translationese in Swedish novels translated from English. In L. Wollin & H. Lindquist (Eds.), *Translation studies in Scandinavia* (pp. 88-95). Lund: CWK Gleerup.
- Gellerstam, M. (1996). Translations as a source for cross-linguistic studies. In K. Aijmer, B. Altenberg & M. Johansson (Eds.), *Languages in contrast: Papers from a symposium on text-based cross-linguistic studies* (pp. 53-62). Lund: Lund University Press.
- Gergen, J. (1994). *Realities and relationships*. Cambridge, Mass: Harvard University Press.

- Gergen, J. (1999). *An invitation to social construction*. London: Sage.
- Goddard, C. (1990). The lexical semantics of good feelings in Yankunytjatjara. *Australian Journal of Linguistics*, 10(2), 257-292.
- Goddard, C. (1991). Anger in the western desert: A case study in the cross-cultural semantics of emotion. *Man*, 26(2), 265-279.
- Goddard, C. (1995). 'Cognitive mapping' or 'verbal explication'? Understanding love on the Malay Archipelago *Semiotica*, 106(3/4), 323-354.
- Goddard, C. (1996). The 'social emotons' of Malay (Bahasa Melayu). *Ethos*, 24(3), 426-464.
- Goddard, C. (1997). Contrastive semantics and cultural psychology: 'Surprise' in Malay and English. *Culture and Psychology*, 3(2), 153-181.
- Goh, C.-L., Asahara, M., & Matsumoto, Y. (2005). Chinese word segmentation by classification of characters. *Computational linguistics and Chinese Language Processing*, 10(3), 381-396.
- Gong, Q. (1997). *Zhongguo yufaxue shi (The history of Chinese syntax)*. Beijing: Yuwen Chubanshe.
- González, M. d. I. Á. G., Mackenzie, J. L., & Álvarez, E. M. G. (Eds.). (2008). *Current trends in contrastive linguistics: Functional and cognitive perspectives*. Amsterdam; Philadelphia: Benjamins.
- Harkins, J. (1994). *Bridging two worlds*. St Lucia: Queensland University Press.
- Harkins, J. (1996). Linguistic and cultural differences in concepts of shame. In D. Parker, R. Dalziell & I. Wright (Eds.), *Shame and the modern self* (pp. 84-96). Melbourne: Australian Scholarly Publishing.
- Harre, R. (1986). An outline of the social constructionist viewpoint. In R. Harre (Ed.), *The social construction of emotions* (pp. 2-14). Oxford: Blackwell.
- Harre, R. (1993). Universals yet again: A test of the 'Wierzbicka thesis'. *Linguistic Sciences*, 15(3), 231-238.
- Harre, R., & Finlay-Jones, R. (1986). Emotion talk across times. In R. Harre (Ed.), *The social construction of emotions* (pp. 220-233). Oxford: Basil Blackwell.
- Hoey, M. (2003). Lexical priming and the properties of text. from <http://www.monabaker.com/tsresources/LexicalPrimingandthePropertiesofText.htm>
- Hoey, M. (2004). Lexical priming and the properties of text. In A. Partington, J. Morley & L. Harrman (Eds.), *Corpora and discourse* (pp. 385-412). Bern: Peter Lang.
- Hoey, M. (2005). *Lexical priming: A new theory of words and language*. London: Routledge.
- Hu, M. (1996). *Cilei wenti kaocha (An investigation of Chinese word classes)*. Beijing: Beijing Yuyan Xueyuan Chubanshe.
- Hu, Y. (1995). *Modern Chinese. Revised Edition*. (Revised ed.). Shanghai: Shanghai Jiaoyu Chubanshe.
- Huang, Z., Harper, M. P., & Wang, W. (2007). *Mandarin part-of-speech tagging and discriminative reranking*. Paper presented at the Empirical Methods on Natural Language Processing Retrieved from <http://www.aclweb.org/anthology/D/D07/D07-1117.pdf>
- Hunston, S. (2002). *Corpora in applied linguistics*. Cambridge: Cambridge University Press.
- Hunston, S. (2007). Semantic prosody revisited. *International Journal of Corpus Linguistics*, 12(2), 249-268.

- Hunston, S., & Francis, G. (1998). Verbs observed: A corpus-driven pedagogic grammar. *Applied Linguistics*, 19(1), 45-72.
- Hunston, S., & Francis, G. (1999). *Pattern grammar: A corpus-driven approach to the lexical grammar of English*. Amsterdam: Benjamins.
- Ivir, V. (1987). Functionalism in contrastive analysis and translation studies. In R. Dirven & V. Fried (Eds.), *Functionalism in linguistics* (pp. 471-481). Amsterdam; Philadelphia: Benjamins.
- Izard, C. E. (1971). *The face of emotion*. New York: Appleton-Century-Crofts.
- Izard, C. E. (1977). *Human emotions*. New York: Plenum Press.
- Izard, C. E. (1991). *The psychology of emotions*. New York: Plenum Press.
- James, C. (1980). *Contrastive analysis*. London: Longman.
- James, W. (1890). *The principles of psychology*. London: Macmillan.
- Jiang, S. (2004). A semantic study of the classifier *dao*. In M. Achard & S. Kemmer (Eds.), *Language, culture and mind* (pp. 429-443). Stanford: CSLI Publications.
- Johansson, S., & Hofland, K. (1994). Towards an English-Norwegian parallel corpus. In U. Fries, G. Tottie & P. Schneider (Eds.), *Creating and using English language corpora* (pp. 25-37). Amsterdam: Rodopi.
- Johansson, S., & Oksefjell, S. (1998). *Corpora and cross-linguistic research*. Amsterdam: Rodopi.
- Josephs, I. E. (1995). The problem of emotions from the perspective of psychological semantics. *Culture and Psychology*, 1(2), 279-288.
- Kennedy, G. (1998). *An introduction to corpus linguistics*. London: Longman.
- Kilgarriff, A. (2009). Putting the corpus into the dictionary. In V. B. Y. Ooi, A. Pakir, I. S. Talib & P. K. W. Tan (Eds.), *Perspectives in lexicography: Asia and beyond* (pp. 239-247). Israel: K Dictionaries Ltd.
- Kilgarriff, A., & Rundell, M. (2002). *Lexical profiling software and its lexicographic applications-a case study*. Paper presented at the EURALEX, Copenhagen.
- Kilgarriff, A., Rychly, S., Smrz, J., & Tugwell, D. (Eds.). (2004). *The Sketch Engine*. In *Proceedings of Euralex. Lorient, France, July: 105-116*.
- Kilgarriff, A., & Tugwell, D. (2001). *WORD SKETCH: Extraction and display of significant collocations for lexicography*. Paper presented at the ACL workshop on COLLOCATION: Computational Extraction, Analysis and Exploitation, Toulouse.
- Kleinman, A. (1980). *Patients and healers in the context of culture*. Berkeley; Los Angeles; London: University of California Press.
- Kondo, F. (2004). *Words denoting emotions: The case of unhappiness in a Japanese-English parallel corpus*. Unpublished manuscript.
- Kövecses, Z. (2000). *Metaphor and emotion: Language, culture, and body in human feeling*. Cambridge: Cambridge University press.
- Krek, S., & Kilgarriff, A. (2006). *Slovene word sketches*. Paper presented at the 5th Slovenian/First International Language Technology Conference, Ljubljana, Slovenia.
- Krzyszowski, T. P. (1990). *Contrasting languages: The scope of contrastive linguistics*. Berlin; New York: M. de Gruyter.
- Lakoff, G. (1987). *Women, fire, and dangerous things: What categories reveal about the mind*. Chicago University of Chicago Press.

- Lakoff, G., & Johnson, M. (1980). *Metaphors we live by*. Chicago: UCP.
- Lakoff, G., & Johnson, M. (1999). *Philosophy in the flesh*. New York: Basic Books.
- Lan, C. (1999). A study of spatial orientational metaphors in Chinese: A cognitive linguistic perspective. *Foreign Languages Teaching and Research*(4), 7-15.
- Lan, C. (2002). A cognitive approach to Up/Down metaphors in English and *Shang/Xia* metaphors in Chinese. In B. Altenberg & G. S. (Eds.), *Lexis in Contrast: Corpus-based approaches* (pp. 161-184): Philadelphia, PA, USA: Benjamins.
- Langacker, R. W. (1987). *Foundations of cognitive grammar*. Stanford: Stanford University Press.
- Langacker, R. W. (1999). Assessing the cognitive linguistic enterprise. In T. Janssen & G. Redeker (Eds.), *Cognitive linguistics, foundations, scope, and methodology* (pp. 13-59). Berlin; New York: M. de Gruyter.
- Lave, J., & Wenger, E. (1991). *Situated learning: Legitimate peripheral participation*. Cambridge: Cambridge University Press.
- Leech, G. (1997). Introducing corpus annotation. In R. Garside, G. Leech & A. McEnery (Eds.), *Corpus annotation: Linguistic information from computer text corpora* (pp. 1-18). London Longman.
- Levy, R. I. (1973). *Tahitians: Mind and experience in the Society Islands*. Chicago: University of Chicago Press.
- Li, C., & Thompson, S. (1981). *Mandarin Chinese: A functional reference grammar*. Berkeley: University of California Press.
- Li, J. (1924). *Xin zhu guoyu wenfa (A new grammar of Chinese)*. Beijing: Shangwu Yinshu Guan.
- Liu, M. (2002). *Mandarin verbal semantics: A corpus-based approach* (2nd ed.). Taipei: Crane Publishing Co.
- Liu, M., Huang, C., Lee, C., & Lee, C.-Y. (2000). When endpoint meets endpoint: A corpus-based lexical semantic study of Mandarin verbs of *throwing*. *Computational linguistics and Chinese Language Processing*, 5(1), 81-96.
- Louw, B. (1993). Irony in the text or insincerity in the writer? The diagnostic potential of semantic prosodies. In M. Baker, G. Francis & E. Tognini-Bonelli (Eds.), *Text and technology* (pp. 157-176). Amsterdam: Benjamins.
- Louw, B. (2000). Contextual prosodic theory: Bring semantic prosodies to life. In C. Heffer, H. Sauntson & G. Fox (Eds.), *Words in context: A tribute to John Sinclair on his retirement*. (pp. 48-94). Birmingham: University of Birmingham.
- Lǔ, S. (1954). Guanyu hanyu cilie de yixie yuanzexing wenti (Some key issues on Chinese word classes). *Journal of Chinese Language*, 9(10), Reprinted in *A collection of papers on Chinese grammar*, pp 256-270 (Shangwu Yinshu Guan, 1984).
- Lǔ, S. (Ed.). (1984). *Hanyu yufa lunwen ji (A collection of papers on Chinese grammar)*. Beijing: Shangwu Yinshu Guan.
- Lutz, C. (1987). Goals, events and understandings in Ifaluk emotion theory. In D. Holland & N. Quinn (Eds.), *Cultural models in language and thought* (pp. 290-312). Cambridge: Cambridge university Press.
- Lutz, C. (1988). *Unnatural remotions: Everyday sentiments on a Micronesian atoll and their challenge to western theory*. Chicago: Chicago University Press.

- Ma, B. (1994). A statistical method to classify word classes. *Journal of Chinese Language* (5), 347-360.
- Machakanja, I. (2008). Conceptual domains and metaphorical mapping in English. *Journal of Language and Communication*(June), 12-32.
- Mahlberg, M. (2005). *English general nouns: A corpus theoretical approach*. Amsterdam: Benjamins.
- Matsumoto. (1989). Cultural inferences on perception of emotion. *Journal of cross-cultural psychology*, 20(1), 92-105.
- Mauranen, A. (2002). Will "translationese" ruin a contrastive study? *Languages in Contrast*, 2(2), 161-186.
- McAndrew, F. T. (1986). A cross-cultural study of recognition thresholds for facial expressions of emotion. *Journal of cross-cultural psychology*, 17(2), 211-224.
- McCawley, J. (1992). Justifying part-of-speech assignment in Mandarin Chinese. *Journal of Chinese Linguistics*, 20(2), 211-246.
- McEnery, T., & Wilson, A. (2001). *Corpus Linguistics. (1st ed.1996)*. Edinburgh: Edinburgh University Press.
- McEnery, T., & Xiao, R. (2002). Domains, text types, aspect marking and English-Chinese translation. *Languages in Contrast*, 2(2), 211-229.
- Mel'c'uk, I. (1988). Semantic description of lexical units in an explanatory combinatorial dictionary: Basic principles and heuristic criteria. *International Journal of Lexicography*, 1(3), 165-188.
- Moon, R. (1998). *Fixed expressions and idioms in English*. Oxford: Clarendon Press.
- Moon, R. (2007). Sinclair, lexicography, and the Cobuild project. *International Journal of Corpus Linguistics*, 12(2), 159-181.
- Moon, R. (2010). What a corpus can tell us about lexis? . In A. O'Keeffe & M. McCarthy (Eds.), *The Routledge handbook of corpus linguistics* (pp. 197-211). London;New York: Routledge.
- Myers, D. G. (1991). *The pursuit of happiness: Who is happy--and why*. New York: W. Morrow.
- O'Halloran, K. A. (2007). Critical discourse analysis and the corpus-informed interpretation of metaphor at the register level. *Applied Linguistics*, 28(1), 1-24.
- Ooi, V. B. Y. (1998). *Computer corpus lexicography*. Edinburgh Edinburgh University Press.
- Ooi, V. B. Y. (2000). 'Language games' on the World Wide Web: Analysing the language of electronic gaming. In C. Heffer & H. Sauntson (Eds.), *Words in Context: A Tribute to John Sinclair on his Retirement* (pp. 110-119). Birmingham: ELR Discourse Analysis Monograph No.18, [CD-ROM publication].
- Ooi, V. B. Y. (2008). The lexis of electronic gaming on the web: A Sinclairian approach. *International Journal of Lexicography*, 21(3), 311-323.
- Packard, J. L. (2000). *The morphology of Chinese: A linguistic and cognitive approach*. Cambridge: Cambridge University Press.
- Partington, A. (2004). "Utterly content in each other's company": semantic prosody and semantic preference. *International Journal of Corpus Linguistics*, 9(1), 131-156.
- Ratnaparkhi, A. (1996). *A maximum entropy model for part-of-speech tagging*. Paper presented at the First Conference on Empirical Methods in Natural Language Processing (EMNLP). Retrieved from <http://acl.ldc.upenn.edu/W/W96/W96-0213.pdf>.

- Renouf, A., & Sinclair, J. (1991). Collocational frameworks in English. In K. Aijmer & B. Altenberg (Eds.), *English corpus linguistics: Studies in honour of Jan Svartvik* (pp. 128-144). London: Longman.
- Rosch, E. H. (1973). On the internal structure of perceptual and semantic categories. In T. E. Moore (Ed.), *Cognitive development and the acquisition of language* (pp. 111-144). New York: Academic Press.
- Russell, J. A. (1991). Culture and the categorization of emotions. *Psychological Bulletin*, 110(3), 426-450.
- Rychly: (2008). A lexicographer-friendly association score. *Proceedings of Recent Advances in Slavonic Natural Language Processing, RASLAN 2008. Brno, pp. 6--9. .*
- Santangelo: (2003). *Sentimental education in Chinese history: An interdisciplinary textual research on Ming and Qing sources*. Leiden: Brill.
- Santos, D. (1998). Perceptions verbs in English and Portuguese. In S. Johansson & S. Oksefjell (Eds.), *Corpora and cross-linguistic research* (pp. 319-342). Amsterdam: Rodopi.
- Sapir, E. (1933). Language. *Encyclopaedia of the Social Sciences (New York)*, 9, 155-169.
- Schachter: (1985). Part of speech systems. In T. Shopen (Ed.), *Language typology and syntactic description, Vol. I: Clause structure* (pp. 3-61). Cambridge: Cambridge University Press.
- Scherer, K. R. (1994). Towards a concept of 'modal emotions'. In P. Ekman & R. J. Davidson (Eds.), *The nature of emotion: Fundamental questions* (pp. 25-31). Oxford; New York: Oxford University Press.
- Schmied, J. (1998). Differences and similarities of close cognates: English *with* and German *mit*. In S. Johansson & S. Oksefjell (Eds.), *Corpora and cross-linguistic research* (pp. 255-276). Amsterdam: Rodopi.
- Scott, M., & Tribble, C. (2006). *Textual patterns: Keyword and corpus analysis in language education*. Amsterdam: Benjamins.
- Sealey, A. (2010). *Researching English language: A resource book for students*. Milton Park, Abingdon, Oxon; New York, NY: Routledge.
- Shen, J. (1999). *Bu duicheng he biaoji lun (On asymmetry and markedness)*. Nanchang: Jiangxi Jiaoyu Chubanshe.
- Shweder, R. (1991). *Thinking through cultures: Expeditions in cultural psychology*. Cambridge: Cambridge University press.
- Shweder, R. (1993). The cultural psychology of the emotions. In M. Lewis & J. M. Haviland (Eds.), *Handbook of emotions* (pp. 417-431).
- Shweder, R. (1994). You're not sick, you're just in love: Emotion as an interpretive system. In P. Ekman & R. J. Davidson (Eds.), *The nature of emotion: Fundamental questions* (pp. 32-44). Oxford; New York: Oxford University Press.
- Sinclair, J. (1985). On the integration of linguistic description. In T. A. v. Dijk (Ed.), *Handbook of Discourse Analysis* (pp. 13-28). London: Academic Press.
- Sinclair, J. (1987). *Looking up: An account of the COBUILD project in lexical computing*. London: Collins.
- Sinclair, J. (1991). *Corpus, concordance, collocation*. Oxford: Oxford University Press.
- Sinclair, J. (1996). The search for units of meaning. *Textus*, IX, 75-106. [Also reprinted in Sinclair, 2004].

- Sinclair, J. (1998). The lexical item. In E. Weigand (Ed.), *Contrastive lexical semantics* (pp. 1-24). Amsterdam; Philadelphia: Benjamins.
- Sinclair, J. (2004). *Trust the text: language, corpus and discourse*. New York: Taylor & Francis.
- Sinclair, J., & Teubert, W. (2004). Interview with John Sinclair, conducted by Wolfgang Teubert. In R. Krishnamurthy, J. Sinclair, S. Jones & R. Daley (Eds.), *English collocation studies: The OSTI report* (pp. xvii-xxix). London: Continuum.
- Stearns, C. Z. (1993). Sadness. In M. Lewis & J. M. Haviland (Eds.), *Handbook of emotions* (pp. 547-561). New York: Guilford.
- Stefanowitsch, A. (2004). Happiness in English and German: A metaphorical-pattern analysis. In M. Achard & S. Kemmer (Eds.), *language, culture and mind* (pp. 137-149). Stanford, Calif.: CSLI Publications.
- Stewart, D. (2010). *Semantic prosody: a critical evaluation*. New York; London: Routledge.
- Stubbs, M. (1996). *Text and corpus analysis: computer assisted studies of language and culture*. Oxford Blackwell.
- Stubbs, M. (1997). Whorf's children: critical comments on critical discourse analysis. In A. Ryan & A. Wray (Eds.), *Evolving models of language* (pp. 100-116). Clevedon: Multilingual Matters.
- Stubbs, M. (2009). The search for units of meaning: Sinclair on empirical semantics. *Applied Linguistics*, 30(115-137).
- Summers, D. (1991). *Longman/Lancaster English language corpus: criteria and design*. Harlow: Longman.
- Sweetser, E. (1990). *From etymology to pragmatics*. Cambridge: Cambridge University Press.
- Tai, J. (1994). Chinese classifier systems and human categorization. In M. Chen & O. Tzeng (Eds.), *Essays in honour of Professor William S-Y Wang* (pp. 479-494). Taipei: Pyramid Publishing Company.
- Tai, J., & Chao, F. (1990). A semantic study of the classifier 'zhang'. *Journal of the Chinese Language Teachers association*, 29(3), 66-78.
- Talmy, L. (1985). Lexicalization patterns: Semantic structure in lexical forms In T. Shopen (Ed.), *Language typology and syntactic description, Vol. III: Grammatical categories and the lexicon* (pp. 57-149). Cambridge: Cambridge University Press.
- Teubert, W. (1996). Comparable or parallel corpora? *International Journal of Lexicography*, 9(3), 238-264.
- Teubert, W. (1999). Corpus linguistics---A partisan view. TELRI Newsletter No.8. (pp. 1-23).
- Teubert, W. (2004a). Language and corpus linguistics. In M. A. K. Halliday, W. Teubert, C. Yallop & A. Cermakova (Eds.), *Lexicology and corpus linguistics: An introduction*. (pp. 73-112). London: Continuum.
- Teubert, W. (2005). My version of corpus linguistics. *International Journal of Corpus Linguistics*, 10(1), 1-13.
- Teubert, W. (2010). *Meaning, discourse and society*. New York: Cambridge University Press.
- Teubert, W. (2004b). When Did We Start Feeling Guilty? In E. Weigand (Ed.), *Emotion in dialogic interaction* (pp. 121-162). Amsterdam: Benjamins.
- Teubert, W., & Cermakova, A. (2004). Directions in corpus Linguistics. In M. A. K. Halliday, W. Teubert, C. Yallop & A. Cermakova (Eds.), *Lexicography and corpus linguistics: An introduction* (pp. 113-165). London: Continuum.

- Thede, S. M., & Harper, M. P. (1999). *A second order hidden Markov model for part-of-speech tagging*. Paper presented at the 37th Annual Meeting of the Association for Computational Linguistics on Computational Linguistics. Retrieved from <http://acl.ldc.upenn.edu/P/P99/P99-1023.pdf>
- Thompson, S. (1988). A discourse approach to the cross-linguistic category "adjective". In J. Hawkins (Ed.), *Explaining language universals*. (pp. 167-210). Oxford: Blackwell.
- Tognini-Bonelli, E. (2001). *Corpus linguistics at work*. Amsterdam: Benjamins.
- Tsai, M., Huang, C., Chen, K., & Ahrens, K. (1998). Towards a representation of verbal semantics: An approach based on near-synonyms. *Computational linguistics and Chinese Language Processing*, 3(1), 62-74.
- Tseng, H., Jurafsky, D., & Manning, C. (2005). *Morphological features help pos tagging of unknown words across language varieties*. Paper presented at the Fourth SIGHAN Workshop on Chinese Language Processing. Retrieved from http://www.stanford.edu/~jurafsky/sighan_pos.pdf.
- Turpin, M. (2002). Body part terms in Kaytetye feeling expressions. *Pragmatics and Cognition*, 10(1/2), 271-305.
- Viberg, A. (1998). Contrasts in polysemy and differentiation: Running and putting in English and Swedish. In S. Johansson & S. Oksefjell (Eds.), *Corpora and cross-linguistic research* (pp. 343-376). Amsterdam: Rodopi.
- Vygotsky, L. S. (1978). *Mind in society: The development of higher psychological processes*. Cambridge Mass: Harvard University Press.
- Wang, J. (2009). *A corpus-based study on the Chinese near-synonymous verbs of running*. Paper presented at the 21st North American Conference on Chinese Linguistics (NACCL-21) (Proceedings. Volume 2. Edited by Yun Xiao, Smithfield, Rhode Island: Bryant University. Pages 399-416).
- Wang, L. (1955). *Zhongguo yufa lilun (A theory of Chinese grammar)*. Beijing: Zhonghua Shuju.
- White, G. M., & Kirkpatrick, J. (1985). *Person, self and experience*. Berkeley: University of California Press.
- Whitsitt, S. (2005). A critique of the concept of semantic prosody. *International Journal of Corpus Linguistics*, 10(3), 283-305.
- Whorf, B. (1956). *Language, thought, and reality: Selected writings of Benjamin Lee Whorf*. Ed. J.B. Carroll. Cambridge: MIT Press.
- Wierzbicka, A. (1972). *Semantic primitives*. Frankfurt: Athenäum.
- Wierzbicka, A. (1986). Human emotions: Universal or culture-specific? *American Anthropologist*, 88(3), 584-594.
- Wierzbicka, A. (1992a). Defining emotion concepts. *Cognitive Science*, 16(4), 539-581.
- Wierzbicka, A. (1992b). *Semantics, culture, and cognition: universal human concepts in culture-specific configurations*. New York: Oxford University Press.
- Wierzbicka, A. (1995). Emotion and facial expression: A semantic perspective. *Culture and Psychology*, 1, 227-258.
- Wierzbicka, A. (1996). *Semantics: primes and universals*. Oxford; New York: Oxford University Press.

- Wierzbicka, A. (1999). *Emotions across languages and cultures: diversity and universals*. Cambridge [England]; New York; Paris: Cambridge University Press; Editions de la Maison des sciences de l'homme.
- Williams, G. C. (1998). Collocational networks: Interlocking patterns of lexis in a corpus of plant biology research articles. *International Journal of Corpus Linguistics*, 3(1), 151-171.
- Williams, G. C. (2006). Book review: Michael Hoey. *Lexical priming: A new theory of words and language*. London: Routledge. 2005. *International Journal of Lexicography*, 19(3), 327-335.
- Xia, F. (2000). The part-of-speech tagging guidelines for the Penn Chinese Treebank (3.0). Retrieved Aug 6, 2010, from <http://www.cis.upenn.edu/~chinese/posguide.3rd.ch.pdf>.
- Xiao, R., & McEnery, T. (2006). Collocation, semantic prosody, and near synonymy: a cross-linguistic perspective. *Applied Linguistics*, 27(1), 103-129.
- Xiao, R., McEnery, T., & Qian, Y. (2006). Passive constructions in English and Chinese: A corpus-based contrastive study. *Languages in Contrast*, 6(1), 109-149.
- Xu, Y. (2002). *Contrastive linguistics*. Shanghai: Shanghai waiyu jiaoyu chubanshe.
- Ye, Z. (2001). An inquiry into 'sadness' in Chinese. In J. Harkins & A. Wierzbicka (Eds.), *Emotions in crosslinguistic perspective* (pp. 359-404). Berlin; New York: M de Gruyter.
- Ye, Z. (2002). Different modes of describing emotions in Chinese: Bodily changes, sensations, and bodily images. *Pragmatics and Cognition*, 10(1/2), 307-339.
- Yip, (2000). *The Chinese lexicon: A comprehensive survey*. London/New York: Routledge.
- Yu, N. (2002). Body and emotion: Body parts in Chinese expression of emotion. *Pragmatics and Cognition*, 10(1/2), 341-367.
- Yu, N. (2003). Metaphor, body, and culture: The Chinese understanding of gallbladder and courage. *Metaphor and Symbol* 18, 13-31.
- Yu, N. (2009). *The Chinese heart in a cognitive perspective: culture, body, and language*. Berlin; New York: Mouton de Gruyter.
- Zhang, B., & Fang, M. (1996). *Hanyu gongneng yufa yanjiu (A study on Chinese functional grammar)*. Nanchang: Jiangxi Jiaoyu Chubanshe.
- Zhang, R. (2009). An interview with Professor Wolfgang Teubert. *Journal of Foreign Languages*, 32(2), 84-93.
- Zhang, R., & Ooi, V. B. Y. (2008). *A corpus-based analysis of qing (情): A contrastive-semantic perspective*. Paper presented at the The International Symposium on Using Corpora in Contrastive and Translation Studies, Hangzhou, China.
- Zhang, Z. (Ed.). (1979). *Hanyu Zhishi (On Chinese language)*. Beijing: Renmin Jiaoyu Chubanshe.
- Zhu, D. (1980). *Xiandai hanyu yufa yanjiu (A study on modern Chinese grammar)*. Beijing: Shangwu Yinshu Guan.
- Zhu, D. (1982). *Yufa jiangyi (Lecture notes on grammar)*. Beijing: Shangwu Yinshu Guan.
- Zhu, D. (1985). *Yufa dawen (Some questions on grammar)*. Beijing: Shangwu Yinshu Guan.

APPENDIX: English glossing of Chinese words

	English Glossing	Chinese Pinyin	Chinese Characters
Extro- version	facial expression	<i>biaoqing/shenqing/ shense/lianse</i>	表情/神情/神色/脸色
	on the face/face/full face	<i>lianshang/lian/manlian</i>	脸上/脸/满脸
	look/see	<i>kan/ kandao</i>	看/看到
	appearance	<i>yangzi</i>	样子
	appear	<i>xiande</i>	显得
	eye	<i>yanjing</i>	眼睛
	expression in one's eyes	<i>yanshen/muguang</i>	眼神/目光
	cry/weep	<i>ku/kuqi</i>	哭/哭泣
	voice	<i>shengyin</i>	声音
	groan	<i>shenyin</i>	呻吟
	tears	<i>yanlei/leishui</i>	眼泪/泪水
	shed tears	<i>luolei/liulei</i>	落泪/流泪
	cry bitterly	<i>tongku</i>	痛哭
	say	<i>shuo/shuodao</i>	说/说道
	express (deeply)	<i>(shen)biao/biaoda/biaoshi</i>	(深)表/表达/表示
	convey	<i>zhuanda</i>	转达
	reveal	<i>liulu/louchu</i>	流露/露出
	manifest	<i>biaoxian</i>	表现
	telegram of condolence	<i>yandian/zhidian</i>	唁电/致电
	form	<i>xingshi</i>	形式
	way	<i>fangshi</i>	方式
	fly the flag at half mast	<i>jiang banqi/xia banqi</i>	降半旗/下半旗
Intro- version	feel	<i>gan(dao)/ganjue/ganshou/ gan/jue(de)/(shen)gan/</i>	感(到)/感觉/感受/感/ 觉(得)/(深)感
	inner heart/heart	<i>neixin/xin</i>	内心/心
	in the heart	<i>xinzhong/xinli</i>	心中/心里
	frame of mind	<i>xinqing</i>	心情
	mood	<i>qingxu</i>	情绪
	emotion/feeling	<i>ganqing/qinggan/qing</i>	感情/情感/情
	mental	<i>jingshen</i>	精神
	mind	<i>xinling</i>	心灵
	soul	<i>linghun</i>	灵魂
	think	<i>xiang</i>	想
Verb	cause/make	<i>rang/ ling/ shi/ jiao</i>	让/令/使/叫
	bring about	<i>zaocheng/dailai/dai</i>	造成/带来/带
	induce/give rise to/ lead to/invite	<i>re/nao/yinqi</i>	惹/闹/引起
	occur/emerge	<i>fasheng/chuxian/ chansheng</i>	发生/出现/产生
	be filled with	<i>chongman</i>	充满
	be filled in one's bosom/chest	<i>manhuai</i>	满怀
	be immersed in	<i>chenjin</i>	沉浸
	fall into/get caught into	<i>xianru</i>	陷入
	suffer	<i>zaoshou/shou</i>	遭受/受
	endure	<i>renshou/chengshou/ jingshou</i>	忍受/承受/经受

	(be forced to) control	<i>(qiang)ren</i>	(强)忍
	relieve	<i>jiechu/jietuo/baituo</i>	解除/解脱/摆脱
	lessen	<i>jianqing/jianshao</i>	减轻/减少
Modi- fier	extremely	<i>wanfen/ji/jile/jidu/jiqi/ jiwei/wubi</i>	万分/极/极了/极度/极 其/极为/无比
	totally	<i>wanquan</i>	完全
	endlessly	<i>wuxian/buyi</i>	无限/不已
	unbearable	<i>nanyirensou/bukan</i>	难以忍受/不堪
	exceedingly	<i>guodu</i>	过度
	utmost	<i>moda</i>	莫大
	enormous	<i>juda/jida</i>	巨大/极大
	very	<i>hen/feichang/shifen /laoda/ting/hao/</i>	很/非常/十分/老大/挺/ 好
	deep	<i>shen(shen)/shenqie</i>	深(深)/深切
	really	<i>zhen/shizai</i>	真/实在
	particularly	<i>tebie</i>	特别
	rather	<i>xiangdang</i>	相当
	a little/slightly	<i>(yi)dian/(yi)xie/youdian/ youxie/(lue)dai</i>	(一)点/(一)些/有点/有 些/(略)带
	faint	<i>dandan de</i>	淡淡的
	a minute quantity	<i>si</i>	丝
	small amount	<i>jifen</i>	几分
	severely	<i>yanzhong</i>	严重
	heavy	<i>chenzhong</i>	沉重
	with grief	<i>chentong</i>	沉痛
	sincere	<i>chengzhi</i>	诚挚
	always/all the time	<i>zongshi/yizhi</i>	总是/一直
	often	<i>changchang</i>	常常
	all the day	<i>zhengtian</i>	整天
	lasting for a short time	<i>yizhen</i>	一阵
	for a long time	<i>changqi</i>	长期
	kind/this kind/that kind	<i>zhong/zhezong/nazhong</i>	种/这种/那种
	portion	<i>fen</i>	份
	number of times	<i>ci</i>	次
Feeling	extremely sad	<i>beitongyujue</i>	悲痛欲绝
	grief	<i>beitong</i>	悲痛
	sorrow	<i>beishang/beiai</i>	悲伤/悲哀
	sad	<i>bei</i>	悲
	sad	<i>nanguo</i>	难过
	sad	<i>shangxin</i>	伤心
	immensely sad/ heartbroken	<i>xinsui</i>	心碎
	mourn	<i>aidao</i>	哀悼
	depression	<i>yiyou/yiyuzheng/ youyuzheng</i>	抑郁/抑郁症/忧郁症
	melancholy	<i>youyu</i>	忧郁
	depressed	<i>yayi</i>	压抑
	melancholy	<i>shanggan</i>	伤感
	gloomy	<i>yinyu</i>	阴郁
	gloomy	<i>menmenbule</i>	闷闷不乐
	dejected	<i>jusang/sangqi/ chuitousangqi</i>	沮丧/丧气/垂头丧气
disheartened	<i>xinhuiyileng/huixinsangqi</i>	心灰意冷/灰心丧气	

	low-spirited	<i>qingxu diluo</i>	情绪低落
	despondent	<i>huixin</i>	灰心
	anguish	<i>tongku</i>	痛苦
	low-spirited	<i>xiaochen</i>	消沉
	discouraged	<i>qi'nei</i>	气馁
	frustrated	<i>shiyi</i>	失意
	pessimistic	<i>beiguan</i>	悲观
	sad/sentimental	<i>shanggan/chouchang</i>	伤感/惆怅
	sympathetic	<i>tongqing</i>	同情
	indignation	<i>fennu</i>	愤怒
	grief and indignation	<i>beifen</i>	悲愤
	annoyance/anger	<i>shengqi/qi</i>	生气/气
	agitated/perturbed	<i>fan/fanzao/xinfan</i>	烦/烦躁/心烦
	anxious/worried	<i>jiaolü/zhaoji</i>	焦虑/着急
	dreary	<i>chenmen</i>	沉闷
	worry	<i>youchou/you/chou/lü</i>	忧愁/忧/愁/虑
	woebegone	<i>choumeikulian</i>	愁眉苦脸
	gloomy	<i>kumen</i>	苦闷
	grievance/feel wronged	<i>weiqu</i>	委屈
	uneasiness	<i>bu'an</i>	不安
	vexation	<i>fannao</i>	烦恼
	worried	<i>kunao</i>	苦恼
	fear	<i>kongju/haipa</i>	恐惧
	shock	<i>zhenjing</i>	震惊
	despair	<i>juewang</i>	绝望
	disappointment	<i>shiwang</i>	失望
	loneliness	<i>gudu/jimo/guji</i>	孤独/寂寞/孤寂
	shame	<i>xiukui</i>	羞愧
	regret	<i>houhui</i>	后悔
	love	<i>ai</i>	爱
	happy	<i>xingfu</i>	幸福
	joy	<i>huanle</i>	欢乐
	gaiety	<i>kuaile</i>	快乐
	delight	<i>gaoxing</i>	高兴
	hope	<i>xiwang</i>	希望
People involved	mother	<i>muqin</i>	母亲
	father	<i>fuqin</i>	父亲
	parents	<i>fumu</i>	父母
	family (member)	<i>qinshu/qinren/jiashu</i>	亲属/亲人/家属
	wife	<i>qizi</i>	妻子
	huaband	<i>zhangfu</i>	丈夫
	son	<i>erzi</i>	儿子
	victim	<i>sinanzhe/li'nanzhe/ yuhaiizhe/shouhaiizhe</i>	死难者/罹难者/ 遇害者/受害者
	the whole nation	<i>quanguo</i>	全国
	people	<i>renmin</i>	人民
	compatriot	<i>tongbao</i>	同胞
	government	<i>zhengfu</i>	政府
	president/chairman/ prime minister/leader	<i>zongtong/zhuxi/zongli/ lingdaoren</i>	总统/主席/总理/ 领导人

	mourner	<i>aidaozhe</i>	哀悼者
Cause	because/why/reason	<i>yin(wei)/weishenme/ yuanyin</i>	因(为)/为什么/原因
	news	<i>xiaoxi</i>	消息
	news about the death of a beloved person	<i>e'hao</i>	噩耗
	failure	<i>shibai</i>	失败
	difficulty	<i>kunnan</i>	困难
	setback	<i>cuozhe</i>	挫折
	hardship	<i>jiannan</i>	艰难
	misfortune	<i>buxing</i>	不幸
	loss	<i>shiqu</i>	失去
	disaster	<i>zainan</i>	灾难
	die/death	<i>si/siqu/siwang/shishi/ qushi</i>	死/死去/死亡/逝世/ 去世
	die of disease	<i>bingshi</i>	病逝
	be killed in a disaster	<i>li'nan/zu'nan/yuhai/ shenwang</i>	罹难/遇难/遇害/身亡
	give one's life for/die a martyr's death	<i>xisheng</i>	牺牲
	loss/lose	<i>shiqu</i>	失去
	departure	<i>likai</i>	离开
	think of	<i>xiangqi/xiangdao</i>	想起/想到
	thought	<i>sixiang</i>	思想
Modi- fies	past	<i>guoqu</i>	过去
	memory	<i>huiyi/jiyi</i>	回忆/记忆
	experience	<i>jingli/tiyan/jingyan</i>	经历/体验/经验
	process	<i>guocheng</i>	过程
	thing/event	<i>shi/shiqing/shijian</i>	事/事情/事件
	past events	<i>wangshi</i>	往事
	story	<i>gushi</i>	故事
	work of art	<i>zuopin</i>	作品
	atmosphere	<i>qifen</i>	气氛
	occasion	<i>changhe</i>	场合
	temperament	<i>qizhi</i>	气质
	personality	<i>gexing</i>	个性
	character	<i>xingge</i>	性格
Others	(of sorrow) be blocked	<i>yujie/yuji</i>	郁结/郁积
	silence	<i>chenmo</i>	沉默
	contemplation	<i>chensi</i>	沉思
	calmness	<i>pingjing</i>	平静
	pale	<i>cangbai</i>	苍白
	smile	<i>xiao/weixiao</i>	笑/微笑
	romantic	<i>langman</i>	浪漫
	gentle	<i>wenrou</i>	温柔
	prince	<i>wangzi</i>	王子
	beauty	<i>meili</i>	美丽
	intestine+broken	<i>chang(duan)</i>	肠(断)
	gall bladder+crack	<i>(dan)lie</i>	(胆)裂
	belly	<i>duzi</i>	肚子
	body	<i>routi/shenti</i>	肉体/身体

	torment	<i>zhemo</i>	折磨
	torture	<i>jian'ao</i>	煎熬
	struggle	<i>zhengzha</i>	挣扎
	powerless/helpless	<i>wukenanhe</i>	无可奈何
	sigh	<i>tanxi</i>	叹息
	console	<i>anwei/weiwen</i>	安慰/慰问
	persuade	<i>quan</i>	劝
	give up	<i>fangqi</i>	放弃
	human kind	<i>renlei</i>	人类
	human nature	<i>renxing</i>	人性
	family	<i>jiating</i>	家庭
	work/job	<i>gongzuo</i>	工作
	society	<i>shehui</i>	社会
	reality	<i>xianshi</i>	现实
	history	<i>lishi</i>	历史
	nation	<i>minzu</i>	民族
	life	<i>rensheng/shenghuo/ shengming</i>	人生/生活/生命
	dilemma	<i>maodun</i>	矛盾
	leave	<i>zou</i>	走
	in mourning	<i>shouxiao</i>	守孝
	mourning wreath	<i>huaquan</i>	花圈
	mourning coach	<i>songzangche</i>	送葬车
	mourning clothes	<i>sangfu</i>	丧服
	go into mourning	<i>fusang</i>	服丧
	announce the death	<i>baosang</i>	报丧
	patient	<i>huanzhe/bingren</i>	患者/病人
	contract/suffer from	<i>huan</i>	患
	treatment	<i>zhiliao</i>	治疗
	symptom	<i>zhengzhuang</i>	症状
	disease	<i>jibing/bing/zheng</i>	疾病/病/症
	nerve	<i>shenjing</i>	神经
	insomnia	<i>shimian</i>	失眠
	suicide	<i>zisha</i>	自杀
	thus/so	<i>bian</i>	便
	nostalgic	<i>huaijiu</i>	怀旧
	energy	<i>qi/nengliang</i>	气/能量
	point	<i>dian</i>	点
	desolate	<i>qiliang</i>	凄凉
	weather	<i>tianqi</i>	天气
	shroud	<i>longzhao</i>	笼罩
	oneself	<i>ziji/ziwo</i>	自己/自我
	life/living	<i>rizi</i>	日子
	have a hard time	<i>nanguo (lit)</i>	难过
	chained	<i>shufu</i>	束缚
	what exactly makes one sad	<i>shangxin chu</i>	伤心处
	between	<i>zhijian</i>	之间
	in /in the middle of	<i>zhizhong/zhong</i>	之中/中