

**A CORPUS-BASED INVESTIGATION OF
IDIOMATIC MULTIWORD UNITS**

by

Lynn E. Grant

**A thesis
submitted to the Victoria University of Wellington
in fulfilment of the
requirements for the degree of
Doctor of Philosophy
in Applied Linguistics**

Victoria University of Wellington

2003

ABSTRACT

Idioms – a type of multiword unit (MWU) – are defined as being non-compositional and in general cannot be understood by adding together the meanings of the individual words that comprise the MWU. Because of this, they present a particular challenge to students who speak English as a second- or foreign-language (ESL/EFL). As a teacher of second-language (L2) learners, it is just that challenge which has motivated this study.

Specifically, there were two main aims of the thesis. In order to know how to teach idioms to ESL/EFL learners, we – as language teachers – need to know how to define and explain them. Therefore, the first aim of the study was to either find an English (L1) definition of an idiom which could clearly distinguish one type from another, and an idiom from a non-idiom, or to develop a new definition. Having not found such a definition, a new definition was put forward, dividing MWUs presently known as idioms into three new groups – core idioms, figuratives, and ONCEs (one non-compositional element). The L1 perspective was adopted for the definition as an L2 perspective would involve considerably more variables.

The second aim was to develop a comprehensive list of one of the three new groups – core idioms – and then try to establish frequency, using a corpus search. A number of steps were taken to compile this list, involving an examination of several sources of written and spoken English. The result was that when the criteria established to define a core idiom – being both non-compositional and non-figurative – were strictly applied to the large collection of MWUs presently known as ‘idioms’, the figure was reduced to only 104 MWUs deemed to be either core idioms or ‘borderline figuratives’ and

‘borderline ONCEs’. Next the British National Corpus (BNC), a corpus of 100 million words, was searched for occurrences of these 104 core idioms and borderlines to establish their frequency. The result of the corpus search showed that none of the core idioms occurs frequently enough to get into the most frequent 5,000 words of English. However, as the motivation to do the study was the desire to find a better way to teach idiomatic MWUs, a brief discussion followed with suggestions for the teaching and learning of these idiomatic MWUs. Finally, some methodological implications and suggestions for future research were put forward, looking at further research which would advance the field of second-language acquisition (SLA) related to the learning of idiomatic MWUs.

ACKNOWLEDGEMENTS

I would like to express my sincere gratitude to the following people for assisting me in a variety of ways throughout my study at Victoria University of Wellington:

- First and foremost, my supervisors Professor Paul Nation and Professor Laurie Bauer for their excellent supervision involving generosity with their time, as well as invaluable advice, guidance and encouragement, given with great patience and good humour.
- The Research Committee of the Faculty of Humanities and Social Sciences, Victoria University of Wellington, for supporting me with Grants.
- Fellow PhD colleagues at Victoria University of Wellington who have provided moral support and lively debate, and helped by finding resources, books, articles – in particular, John Macalister, Dianne Bardsley, and Nany Kurnia – as well as Administrative staff who arranged office space during my many visits to Wellington.
- Professor Graeme Kennedy for his advice on corpora.
- My Wellington ‘family’ – Philippa, Bill, Morgan and Grace – who accommodated and fed me during my numerous trips to Wellington.
- Colleagues in my workplace, the School of Languages at the Auckland University of Technology, for their interest and support. My thanks in particular to my Head of School, Associate Professor Ron Holt, for allowing the flexibility of teaching hours, the financial support which allowed research time out of the classroom, and help with technical equipment and programmes.
- Friends, colleagues, and neighbours – in particular Rosemary Wette, Darryl Hocking, Gay and Tim Devlin, Brian Paltridge, Jude Buckler and Graeme

Mackay, Alison and Hugo Hoffmann, Jude Day and Ron Griffin, Andrew, Donald and Stuart Mackay, Clare Conway, Claudia and Callum Danby-Scott, Michael Harland, Lily Lim, Sue Knox – who acted as ‘inter-rater reliability testers’, collected idioms from newspapers and from teenagers, donated books on idioms, provided invaluable technical support with the computer, as well as moral support and encouragement.

- And last but not least, my family, in particular my daughter Anthea for her patience and understanding, and my sisters, Merle Ferguson and Brenda Grant, for their moral and financial support.

It is the support of all of the above that has enabled the completion of this study, and for this I remain eternally grateful.

TABLE OF CONTENTS

CHAPTER 1 INTRODUCTION	10
1.1 BACKGROUND	10
1.2 FIRST STAGE OF THE STUDY.....	12
1.3 SECOND STAGE OF THE STUDY.....	13
CHAPTER 2 DEFINING IDIOMS	15
2.1 INTRODUCTION	15
2.2 WORD COMBINATIONS.....	15
2.2.1 Collocations.....	16
2.2.2 Collocations vs idioms.....	18
2.3 IDIOMS AND IDIOMATICITY	19
2.4 TYPES OF IDIOMS	22
2.5 CLASSIFICATIONS OF IDIOMS.....	25
2.5.1 Semantic classifications of idioms.....	25
2.5.2 Syntactic classifications of idioms	28
2.5.3 Functional classifications of idioms	30
2.6 A SCALE OF IDIOMATICITY	32
2.7 THE LACK OF A CLEAR DEFINITION	34
2.8 IDENTIFYING CRITERIA FOR DEFINING IDIOMS.....	36
2.8.1 Non-compositionality:.....	36
2.8.2 Part compositionality:.....	37
2.8.3 Institutionalisation:	37
2.8.4 Frozenness:	37
2.9 THE ‘COMPOSITIONALITY’ OF LANGUAGE.....	41
2.9.1 Problems with ‘compositionality’	42
2.9.2 The ‘naïve approach’ to compositionality	44
2.10 ESTABLISHING NEW CRITERIA FOR DEFINING IDIOMS.....	46
2.11 FIGURATIVE LANGUAGE	47
2.11.1 Defining ‘figurative’.....	47
2.11.2 Figurative language vs idioms.....	50
2.12 THE NEW DEFINITION OF AN ‘IDIOM’	52
2.13 CONCLUSION.....	53
CHAPTER 3 DEVELOPING AN ‘IDIOM TEST’	55
3.1 INTRODUCTION	55
3.2 THE CRITERIA FOR DEFINING AN ‘IDIOM’	55
3.3 THE ‘IDIOM TEST’	56
3.3.1 Applying the test to MWUs presently classified as idioms.....	57
3.3.2 Issues to take into consideration when applying the ‘test’	58
3.4 TESTING THE ‘IDIOM TEST’	59
3.4.1 Test 1	60
3.4.2 Test 2	63
3.4.3 Results	65
3.5 CONCLUSION.....	66
CHAPTER 4 DEVELOPING A ‘CORE IDIOM’ LIST FOR THE CORPUS SEARCH.....	68
4.1 INTRODUCTION	68
4.1.1 Applying the core idiom criteria.....	68
4.2 CHOOSING AN IDIOM DICTIONARY	70
4.3 ESTABLISHING THE ACCURACY OF ‘IDIOM’ NUMBERS.....	71

4.4	BORDERLINE CATEGORY.....	73
4.5	CONDUCTING AN ‘INTER-RATER RELIABILITY CHECK’	74
4.5.1	Interpreting idioms	75
4.6	THE CORE IDIOM AND BORDERLINE LIST	77
4.7	ADDING TO THE CORE IDIOMS LIST – OTHER IDIOM DICTIONARIES	78
4.7.1	Core idioms found in other idiom dictionaries.....	79
4.7.2	Hyphenated and non-hyphenated MWUs.....	80
4.8	WIDENING THE SEARCH.....	81
4.9	ADDITIONAL SOURCES OF CORE IDIOMS.....	82
4.9.1	English teaching course books	82
4.9.2	Vocabulary teaching books	83
4.9.3	Journal articles and book chapters written about idioms.....	84
4.9.4	Horoscopes	84
4.9.5	Sports reports.....	85
4.9.6	Scripted conversational television programmes	86
4.9.7	Unscripted conversational television programmes.....	87
4.10	THE FINAL CORE IDIOM LIST	89
4.10	MWUs FROM <i>CAMBRIDGE</i> NOT MEETING THE CORE IDIOM CRITERIA.....	90
4.12	ADDITIONAL MWUs NOT MEETING THE ‘CORE IDIOMS’ CRITERIA	93
4.13	CONCLUSION	95
CHAPTER 5 CONDUCTING THE CORPUS SEARCH AND RESULTS OF THE CORPUS SEARCH.....		97
5.1	INTRODUCTION OF THE CORPUS SEARCH	97
5.2	USING THE BNC SEARCH ENGINE, ‘SARA’	98
5.2.1	Conducting the initial search	99
5.2.2	Conducting a second search	100
5.3	POTENTIAL PROBLEMS CONDUCTION A SEARCH IN THE BNC... ..	101
5.4	MAKING A JUDGEMENT CALL	102
5.4.1	Identical sentences.....	102
5.4.2	Core idioms in isolation.....	102
5.4.3	Literal vs idiomatic.....	102
5.4.4	Different spellings	103
5.5	CONCLUSION OF THE CORPUS SEARCH.....	103
5.6	THE CORPUS SEARCH RESULTS	104
5.7	DECIDING HOW TO LIST THE RESULTS OF THE CORPUS SEARCH	106
5.7.1	How idiom dictionaries list their ‘idioms’	106
5.8	FORMAT FOR RESULTS	109
5.8.1	Examples of the ‘Format for Results’	110
5.8.2	Relevance of the ‘criteria’ to teachers and learners.....	113
5.8.3	Relevance of ‘Format for Results’ to teachers and learners.....	114
5.9	OCCURRENCES INVOLVING VARIATIONS.....	116
5.9.1	Variations in hyphenation.....	116
5.9.2	Variations in spelling.....	118
5.9.3	Variations in singulars/plurals	118
5.9.4	Truncated core idioms	120
5.9.5	Variations in lexical form.....	120
5.9.6	Frozen core idioms	123
5.10	NUMBER OF VARIATIONS OF CORE IDIOMS	124

5.11	COLLOCATIONS AND CORE IDIOMS	126
5.11.1	Noun core idiom collocations.....	127
5.11.2	Verb core idiom collocations.....	128
5.11.3	Adjective core idiom collocations	130
5.11.4	Adverb core idiom collocations.....	131
5.12	DISTINGUISHING BETWEEN CORE IDIOMS AND LITERAL MWUs	132
5.13	RELEVANCE OF THE INFREQUENCY OF CORE IDIOMS.....	133
5.13.1	Top 20 core idioms from the BNC compared to the WC / WSC.....	138
5.14	FREQUENCY OF FIGURATIVES AND ONCEs	139
5.15	LISTING THE RESULTS IN APPENDICES.....	142
5.16	APPLYING SECOND LANGUAGE ACQUISITION (SLA) THEORY TO LEARNING IDIOMATIC MWUs.....	144
5.17	LEARNING CORE IDIOMS	146
5.18	LEARNING FIGURATIVES AND ONCEs.....	149
CHAPTER 6 SUMMARY OF CONCLUSIONS AND IMPLICATIONS FOR FURTHER RESEARCH		159
6.1	SUMMARY	159
6.2	IMPLICATIONS	165
APPENDICES		172
APPENDIX 1: The 20 most frequent core idioms.....		172
APPENDIX 2: Core idioms occurring 31-47 times.....		172
APPENDIX 3: Core idioms occurring 21-30 times.....		173
APPENDIX 4: Core idioms occurring 11-20 times.....		173
APPENDIX 5: Core idioms occurring 10 or fewer times.....		174
APPENDIX 6: Core idioms with zero occurrences in the BNC.....		174
APPENDIX 7: Complete alphabetical list of core idioms and borderlines		175
APPENDIX 8: 'Most frequent idioms' in both <i>Cambridge</i> & COBUILD.....		176
APPENDIX 9: Inter-rater reliability test 1		177
APPENDIX 10: Inter-rater reliability test 2.....		184
APPENDIX 11: 'Idiom' articles and book chapters.....		188
BIBLIOGRAPHY		195

TABLE OF TABLES AND FIGURES

Table 1 Semantic classification of idioms	27
Table 2 Fernando and Flavell's examples of transparency and opacity	34
Table 3 Idiom dictionaries used as a source of idioms	70
Table 4 The 58 core idioms found in <i>Cambridge</i>	77
Table 5 The 12 borderline figuratives/ONCEs found in <i>Cambridge</i>	78
Table 6 Core idioms added from other idiom dictionaries	79
Table 7 Borderline figuratives/ONCEs added from other idiom dictionaries	79
Table 8 Hyphenated words now included from other dictionaries	81
Table 9 Some English teaching course books popular in Auckland	83
Table 10 Some English teaching vocabulary books popular in Auckland	83
Table 11 Sources and dates of horoscopes included	85
Table 12 Newspaper sources of sport reports	85
Table 13 Scripted television programmes used for collecting idiomatic MWUs	86
Table 14 Unscripted television programmes used for collecting idiomatic MWUs	88
Table 15 Core idioms and borderlines added to original list from other sources	89
Table 16 Some <i>Cambridge</i> MWUs containing words not found elsewhere, so not meeting the core idioms criteria	90
Table 17 Some <i>Cambridge</i> foreign word MWUs not meeting the core idioms criteria	90
Table 18 Some <i>Cambridge</i> hyphenated words not meeting the core idiom criteria	91
Table 19 Some <i>Cambridge</i> personal name and fictional character MWUs not meeting the core idiom criteria	92
Table 20 Some <i>Cambridge</i> literal MWUs not meeting the core idiom criteria	92
Table 21 Some <i>Cambridge</i> euphemism MWUs not meeting the core idiom criteria	92
Table 22 Some literal and euphemism MWUs not meeting the core idiom criteria	93
Table 23 Some strengthener MWUs not meeting the core idiom criteria	94
Table 24 Some truncated idiom dictionary MWUs not meeting the core idiom criteria	94
Table 25 Some expressions of emotion not meeting the core idiom criteria	94
Table 26 Some literary, cultural, biblical or historical MWUs not meeting the core idiom criteria	95
Table 27 Some rhyming slang MWUs not meeting the core idiom criteria	95
Table 28 Potentially difficult search results	101
Table 29 Most to least frequent core idioms and number of BNC occurrences	105
Table 30 How idiom dictionaries list idiomatic MWUs	107
Table 31 How core idioms are listed on the attached CD	108
Table 32 Relevance of 'Format for Results' to teacher and/or learners	115
Table 33 A sample of occurrences of hyphenated and non-hyphenated variations	117
Table 34 Unexpected variations involving hyphenation found in the BNC	117
Table 35 Range of differences in spelling of core idioms	118
Table 36 Range of core idioms allowing or not allowing the plural form	119
Table 37 Variations in individual core idioms or borderlines	122
Table 38 Frozen core idioms	124
Table 39 Number and percentage of core idioms / borderlines which allow variations	125
Table 40 Noun core idioms which collocate mainly with the verb 'to be'	127
Table 41 Nouns/pronouns which collocate with (<i>go/goes/went/gone</i>) by the board(s)	128
Table 42 Nouns/pronouns which collocate with four core idioms beginning with a verb	128

Table 43 Verbs which collocate with <i>by and by</i> and <i>by and large</i>	129
Table 44 Verbs which collocate with <i>out of hand</i>	129
Table 45 Nouns which collocate with <i>out and out</i>	130
Table 46 Descriptors which collocate with <i>red herring</i> and <i>white elephant</i>	130
Table 47 Internal modification of <i>cut __ ice with, cut __ dash, and come __ cropper</i>	131
Table 48 Verbs which collocate with <i>like the clappers, to a T</i> and <i>like nobody's</i> <i>business</i>	131
Table 49 Core idioms and borderlines where literal equivalents are not possible.....	134
Table 50 Number of occurrences of core idioms and their literal equivalents in the BNC	135
Table 51 Examples of BNC core idiom frequency in written and spoken English ...	137
Table 52 The 20 most frequent core idioms in the BNC compared to the WC / WSC	138
Table 53 BNC occurrences of 'idioms' marked as frequent in <i>Cambridge</i>	141
Table 54 Frequency of occurrence and number / ratio of core idiom occurrences.....	143
Table 55 Examples of truncated proverbs found in the BNC	152
Table 56 Examples of BNC core idioms numbering '10 or fewer' and '20 or fewer'	163
 Figure 1 Example of a 'visual idiom' of <i>new broom</i>	153

CHAPTER 1

INTRODUCTION

1.1 BACKGROUND

Idioms, a type of multi-word unit (MWU), are not well-defined. That is, the term ‘idiom’ is applied to a wide range of expressions that do not admit a single definition. As a result, a diverse group of multiword units are called idioms, and these cannot be treated uniformly as part of the teaching/learning task, making them difficult for language teachers to explain and teach to learners. Learners can check a dictionary for the meaning of a particular idiom, but gain no real understanding of why it is listed in this particular dictionary or what it has in common with other entries. The MWU entries range from those at the literal/transparent end of the continuum (*Act your age!*, *It’s a free country*, *Seeing is believing*), to those at the opposite/opaque end (*pull someone’s leg*, *red herring*, *kick the bucket*), from those better known as ‘fixed expressions’ (*be that as it may*, *far be it from me to*, *as and when*), to similes (*as old as the hills*, *like a fish out of water*, *as hard as nails*) and proverbs (*The early bird catches the worm*, *An apple a day keeps the doctor away*), from figurative language requiring little explanation (*get the green light*, *be skating on thin ice*, *add fuel to the fire*) to expressions of surprise requiring more (*For crying out loud!*, *Stone the crows!*, *Hell’s bells!*), from drinking toasts (*Here’s mud in your eye*, *Down the hatch!*), to ironic expressions (*Tell me about it!*, *a likely story*, *you don’t have to be a rocket scientist*) and ones whose meaning can be affected by the tone used (*Just my luck*, *big deal*, *Do me a favour!*), from Biblical/historical ones (*the prodigal son*, *turn the other cheek*, *turn*

swords into ploughshares) to ones requiring a knowledge of the literature/culture (*a Jekyll and Hyde*, *Heath Robinson*, *Mickey Mouse*), from sexual (*a bit of how's your father*, *rumpy-pumpy*, *rough trade*) and vulgar/taboo ones (*lick my arse!*, *shove something up your arse*, *ass-kisser*) to foreign ones (*al fresco*, *a priori*, *carte blanche*). Although all of the above MWUs are from the same idiom dictionary (*Cambridge International Dictionary of Idioms*, McCarthy & Walter, 1998), it is difficult to explain to an ESL/EFL learner what this diverse range of MWUs has in common, or to give a simple explanation applicable to all of them.

Over the past six decades, much work has been done on idioms, with different linguists suggesting a variety of definitions. Almost sixty years ago, Roberts (1944, p.300), noting that 'idiom' comes from the Greek *idios*, for 'one's own, peculiar, private, strange', described idiom as "the idiosyncrasy of permutation which a given language exhibits in contradistinction to all languages or a given period exhibits in contrast to all previous periods". Roberts (1944), while describing both the generally idiomatic as well as figures of speech, also noted another element of the definition: "a group of words carrying as a whole a meaning not to be gathered from its component parts separately considered". In the next two decades, this element of compositionality gained importance in the description of idioms proper, with the meaning of an idiom being different from the compositional function of its parts (Hockett, 1958; Weinreich, 1969). This element of compositionality has remained as one of the most common elements in any definition of an idiom.

While linguists and language teachers alike have referred to the importance of teaching idioms to ESL/EFL students, especially those idioms in frequent use, no one has yet established a list of idioms based on frequency. In addition, no one has attempted to apply the theory of Second Language Acquisition (SLA), in particular how it related to vocabulary learning, to learning idioms. These concerns can be expressed in the following research questions:

1. How can we more clearly define 'idioms' so that we can divide the present large collection of MWUs presently known as idioms into separate categories with distinct criteria?
2. How can we establish frequency of these MWUs in order to know which ones to begin to teach our ESL/EFL learners?
3. How can we apply the theories of SLA, in particular as it relates to vocabulary learning, to the teaching, learning and acquisition of idiomatic MWUs?

It is matters such as these which have prompted this particular study, which has two main stages.

1.2 FIRST STAGE OF THE STUDY

The two stages of this study are related to MWUs known as idioms. Because idioms are not well-defined, they are difficult for language teachers to explain to learners, and because they cannot be understood by adding together the meanings of the individual words, they are confusing to ESL/EFL learners. To add to this confusion, they also vary in form from compound words (*willy-nilly*, *elbow grease*), to verb and prepositional phrases (*kick the bucket*, *in the pipeline*), to clauses (*go like a rocket*), to full sentences (*If Mohammed will not*

go to the mountain, the mountain must come to Mohammed). They therefore present significant difficulties to both the language teacher – in knowing how best to teach them and which ones to teach first – and the language learners – in knowing how to first interpret them and then use them accurately and appropriately (Alexander, 1987; Cooper, 1998, 1999; Cowie, 1988; Howarth, 1998; Irujo, 1986; McCarthy, 1998).

There are seven chapters comprising the thesis. Chapter 2 begins by reviewing the literature to find out whether any previous definition can clearly distinguish one type of idiom from another, as well as distinguish an idiom from a non-idiom. If not, the aim was to come up with a definition of an idiom that would make this distinction. As no clear definition was found, the study suggests a definition that divides idioms into three different categories.

Having developed a new definition, Chapter 3 then looks at constructing a test that could be reliably applied to a group of idioms, such as those presently found in idiom dictionaries. The test divides the idioms into the three newly established categories – figuratives, ONCEs (one non-compositional element), and core idioms, plus a ‘borderline’ category. This completes the first stage of the study.

1.3 SECOND STAGE OF THE STUDY

The second stage of the study consists of two parts: compiling a comprehensive list of one of the three new categories – that of core idioms (CIs) – and doing a corpus search of them to establish their frequency. As it

was not possible to compile a list and conduct a search for all three of the categories, core idioms was the category of idioms chosen as the one likely to cause the most difficulty for language learners because of their non-compositionality and the lack of clues to understanding their meaning. Therefore, Chapter 4 describes the steps involved in developing as comprehensive a list of core idioms as possible. This begins by surveying a number of idiom dictionaries on the market and choosing one as the initial source of core idioms, and then examining not only other idiom dictionaries but also a number of other sources of both written and spoken English.

Chapter 5 discusses both the corpus search and the results of the search, looking first at how the search was conducted, the search tools used, problems that occurred during the search, and how frequency was established. The results of the corpus search are given, showing that none of the core idioms occurs frequently enough to merit inclusion in any high frequency vocabulary list. As frequency of occurrence is not the only reason for learning vocabulary, the chapter applies Second Language Acquisition (SLA) theory, in particular as it relates to learning vocabulary, to suggest strategies (memorising, noticing, reading, communicative activities, and so on) for the teaching/learning and acquisition of all idiomatic MWUs.

Chapter 6 concludes with a discussion of the implications of the study and some suggestions for future research which would make a contribution to the study of second-language acquisition with regard to MWUs presently known as 'idioms'.

CHAPTER 2

DEFINING IDIOMS

2.1 INTRODUCTION

This chapter examines how word combinations, in particular MWUs known as idioms, have been defined in the past. First it looks at collocations and how they differ from idioms, and distinguishes between the terms ‘idiom’ and ‘idiomaticity’. The chapter then details the various types and classifications of idioms, as well as the previous attempts to develop a scale or continuum of idiomaticity. Next it examines the three main criteria used to define, or describe, idioms: compositionality, institutionalisation, and frozenness/fixedness, with compositionality emerging as the key criterion. A clear definition of compositionality is given. The chapter then examines figurative language and gives a second clear definition of figurativeness. Finally two criteria are identified to give a new definition of what will be known as a core idiom (CI).

2.2 WORD COMBINATIONS

Idioms – in the sense of ‘colourful phrases listed in idiom dictionaries’ – are not well-defined. They are a type of multi-word unit (MWU), so to define idioms, we must first define MWUs. MWUs can be loosely defined as fixed and recurrent patterns of lexical material sanctioned by usage. At one end of this we have collocations.

Palmer (1933, p.5) illustrates the difficulty of finding a clear definition by searching for a word “that distinguishes *To answer the door* from *To answer a question*, *To strike while the iron’s hot* from *To go while the weather’s fine*, *To take the chair* from *To take a chair*, *The next best* from *The next house*, *I think not* from *I don’t think so*, *Common sense* from *Common glass*, or *Thank you* from *I thanked him*.” The word finally settled upon to explain the first of each pair of examples is ‘collocation’.

2.2.1 Collocations

Collocation is a term widely used now to describe “actual words in habitual company” (Firth, 1957, p.14) or words frequently found in combination with each other. Aisenstadt (1979, pp.1-2) and Sinclair (1991a, pp.109-112) point out that collocations may be free/open or restricted in the sense that it may be possible, depending on co-occurrence restrictions, to substitute both or one constituent.

With restricted collocations (RCs):

- a verb may be found in combination with a restricted number of nouns: face the facts/truth/problem/circumstances (Aisenstadt, 1979, pp.79-80), but in these RCs the verb ‘face’ has a special non-literal sense. In other examples of RCs, a noun may combine with a restricted number of adjectives: red/auburn/blond/brown/black/grey/white hair

With free or open collocations:

- the same verb ‘face’, this time with its literal meaning of ‘positioned opposite or looking in that direction’, can combine freely with a large number of nouns: face the front/door/window/sun/teacher, etc. In the same way, a noun such as ‘bag’ can also combine with a large number of adjectives: red/green/black/large/soft/book/cloth, etc..

Exploiting polysemy, which is central to collocations, is something learners must be made aware of. While free collocations are a matter of syntax, it is the restricted collocations that will be relevant to the definition of idioms.

Sinclair (1991a, pp.109-112) discusses restricted and free collocations as the result of two principles which affect the word choices we make as users of a language:

1. *the open-choice principle*, a 'slot-and-filler' way of seeing language text as the result of a very large number of complex choices. For example, 'set the cup on the bench', 'put the mug in the cupboard' and so on.
2. *the idiom principle*, when a language user has available to them a large number of what Sinclair calls 'semi-preconstructed phrases' that constitute single choices, even though they might appear to be analysable into segments. For example, 'of course', 'set eyes on' and so on.

Sinclair notes that collocation, an illustration of the 'principle of idiom', is both more common and more elusive than previously thought and that the two 'principles' of language are in fact incompatible with one another, with the *idiom principle* dominating, and only when unsuccessful is a switch made to the *open choice principle*.

The term 'collocation' is used to cover a variety of word combinations known by a variety of names (Becker, 1975): prefabricated routines, prefabricated patterns, sentence builders, unassimilated fragments, formulaic speech, lexicalized sentence stems, cliches, deictic locutions, idioms, non-canonical forms, polywords, phrasal constraints, situational utterances, verbatim texts, fixed phrases, set phrases.

For the purposes of this research, the preferred term will be ‘multiword units’ (MWUs) and an attempt will be made to develop a theory of one type of MWU – idioms. First we will look at how collocations compare to idioms.

2.2.2 Collocations vs idioms

Kjellmer (1994) considers idioms a type or subset of collocation, while others consider ‘restricted collocations’ (*cardinal error/sin/virtue/grace*) a type of idiom (Cowie & Mackin, 1975; Cowie, Mackin, & McCaig, 1983). Aisenstadt (1979, p.1) argues that collocations differ from idioms as “R[estricted] C[ollocation]s are not idiomatic in meaning; they do not form one semantic unit; their meaning is made up as the sum of the meanings of their constituents. They have a much greater variability and usually occur in patterns with a number of interchangeable constituents”. According to Aisenstadt (1979), while ‘face the facts/truth/problem/circumstances’ are collocations, ‘face the music’ is an idiom. We have already noted that the verb ‘face’ is used with its literal meaning in the collocation and its non-literal sense in the idiom (see 2.2.1). In addition, restrictions regarding collocations are dependent on both the grammar and the lexicon. Aisenstadt (1979) implies that such semi-productive patterns should be handled as part of the grammar. With regard to collocations and idioms, still others claim there is a continuum ranging from collocations at one end to idioms at the other. Kjellmer (1994, p.xxxiv) explains: “The borderland between idioms and other collocations is then a very fuzzy area, since even the least idiomatic of idioms would normally be collocational, nevertheless.” In other words, while collocations and idioms are both MWUs, open collocations are the loosest kind of MWU and are not relevant to the comparison of the two categories. Tighter collocations come

under the heading of what have been called ‘phrasal lexemes’ (Moon, 1998b, p.79): “the whole range of fixed and semi-fixed complex items...that for reasons of semantics, lexico-grammar, or pragmatics are regarded as holistic units”. For the purposes of this study, I consider ‘multiword unit’ to be an umbrella term covering a variety of phrases, with both ‘idioms’ and ‘collocations’ being distinct types. However, like Moon (1998b, p.79), I am “specifically excluding phrasal verbs such as *give up* and *stick out*, compound nouns, adjectives, and verbs such as *civil servant*, *self-raising*, and *freeze-dry*; and foreign phrases such as *en passant* and *caveat emptor*.” The category of compound words is a complex one, however, and we will revisit it later (see 4.11). My reason for excluding phrasal verbs is that although many of them are idiomatic (for a test of these, see Cowie, 1993), they are such a large group of MWUs that they merit separate and thorough research of their own, as do restricted collocations. I now turn to a definition of ‘idiom’ and ‘idiomaticity’.

2.3 IDIOMS AND IDIOMATICITY

As Chafe has argued:

The importance of idioms in any language cannot be doubted. Their ubiquity makes them anything but a marginal phenomenon, and surely a linguistic theory has an obligation to explain them in a natural way. (1968, p.111)

Weinreich further suggests:

Idiomaticity is important for this reason, if for no other, that there is so much of it in every language. (1969, p.23)

'Idiomaticity' is used ambiguously in the literature. It represents both the quality of being marked by idioms, as well as the ability to speak a fluent and appropriate version of a language. The terms 'idiom' or "type one idiomaticity" and 'idiomaticity' or "type two idiomaticity" (Alexander, 1987), although related, are not identical. The second term 'idiomaticity' applies to how a native speaker of a language successfully combines words, or units of language, in a native-like manner, or "a particular manner of expressing something in a language, music, art, and so on, which characterizes a person or group" (Moon, 1998a, p.3). 'Idiomaticity' is also demonstrated by Sinclair's 'idiom principle' (see 2.2.1), giving a language user a choice of using a number of semi-preconstructed phrases rather than stringing together individual words. To reduce the confusion that exists between various uses of 'idiom' and 'idiomaticity', I will replace the latter with 'nativelike' (cf. 'nativelike selection', Pawley & Syder, 1983, pp.191-2). Having both fluent and nativelike control of a language may depend on knowing a body of 'sentence stems' which are 'institutionalized' or 'lexicalized' (Pawley & Syder, 1983). So when a person takes leave of another in New Zealand with 'See you later', the speaker is showing this nativelike use of the language, in the same way that a native English speaker knows to say 'headache' and 'toothache' but not 'armache' or 'eyeache'. Therefore, the term 'nativelike' will refer to 'a variety of word combinations commonly used by native speakers of a language', and 'idiomatic' will henceforth mean 'having to do with idioms'.

In contrast to 'nativelike', the word 'idiom' refers to a much smaller group of MWUs, those 'fixed/set phrases' which allow limited lexical flexibility and whose meanings cannot be understood by adding together the meanings of

their separate constituents. While all idioms illustrate what is known as idiomaticity (Fernando, 1996) or are what I call nativelike, the reverse is not true. In other words, the use of an idiom like *kick the bucket* meaning ‘die’ illustrates nativelike selection of English but saying that someone has died, although it may be perfectly nativelike, does not illustrate the use of an idiom. Because some idiomatic strings have both a literal and a non-literal meaning (kick the bucket: ‘strike the pail with your foot’, *kick the bucket* = ‘die’; bread and butter: ‘two food items’, *bread and butter* = ‘main income-producing activity’; red tape: ‘coloured ribbon’, *red tape* = ‘excessive bureaucracy’), contextual clues may be necessary to distinguish whether the MWU has a literal or idiomatic interpretation. The term ‘literal’ is being used here as “nonmetaphorical literality: directly meaningful language – not language that is understood, even partly, in terms of something else” (Lakoff, 1986, p.292). Moon (1998a) finds the literal meanings of such phrases as above are always rarer than the idiomatic interpretation, while others (Cowie *et al.*, 1983, p.xiii) go so far as to claim that the literal senses of expressions such as *beat one’s breast* and *burn one’s boats* “do not survive alongside their figurative ones in normal, everyday use”.

Although several criteria have been suggested in the literature to define (or more accurately – describe) idioms and various other MWUs, lists of idioms in dictionaries and teaching texts nevertheless provide a heterogeneous set of items under this all-encompassing heading. In the past, the most commonly accepted of the criteria for defining idioms included some or all of these criteria (Fernando & Flavell, 1981, p.17):

1. the meaning of an idiom is *not* the result of the compositional function of its constituents
2. an idiom is a unit that either has a homonymous literal counterpart or at least individual constituents that are literal, though the expression as a whole would not be interpreted literally
3. idioms are transformationally deficient in one way or another
4. idioms constitute set expressions in a given language
5. idioms are institutionalized

Considerable work on idioms has been done by lexicographers (Cowie & Mackin, 1975; Cowie *et al.*, 1983; Sinclair & Moon, 1989, 1995), Eastern European phraseologists (Arnold, 1973; Lipka, 1991; Mel'cuk, 1988; Mel'cuk, Arbatchewsky-Jumarie, Elnitsky, Iordanskaja, & Lessard, 1984), lexical semanticists (Cruse, 1986), lexicologists (Carter, 1987), and researchers into vocabulary in language teaching (Carter & McCarthy, 1988). The criterion used most often to define idioms is the one related to their non-compositionality. In other words, the meaning of idioms cannot be predicted from the meaning of their constituent parts. My aim is to provide a more restrictive definition of idiom, but also to provide a reliable way of distinguishing whether something is or is not an idiom.

2.4 TYPES OF IDIOMS

The lack of a specific definition has led to the term 'idiom' encompassing a large variety of different types of MWU. As early as the 1920s and 1930s, lexicographers such as Palmer (1933) did work on MWUs, in particular collocations, but did not distinguish between the more or less idiomatic collocations. Later Hockett (1958) described a range of types of construction he terms 'idioms' (substitutes, proper names, abbreviations, English phrasal compounds, figures of speech and slang) which seem to reflect nativelike selection rather than idiomaticity. Malkiel's (1959) detailed description of

‘binomials’ (*odds and ends*, *spick and span*) made a useful contribution to the study of both idioms and nativelike selection, while Makkai’s work, starting with his doctoral thesis (Makkai, 1965), and later published as a book (Makkai, 1972) makes a significant contribution to the field. Makkai provides a division of these idiomatic MWUs into a heterogeneous set of phrase and sentence ‘types’, the most common of which include:

- lexemic idioms:
 - a) phrasal verbs: *make up*, *turn out*, *bring up*, etc.
 - b) tournures (‘turn of phrase’): *fly off the handle*, *rain cats and dogs*, *kick the bucket*, *have it out with somebody*, *be well-off*, etc.
 - c) irreversible binomials: *pepper and salt* (sic), *coffee and cream*, etc.
 - d) phrasal compounds: *hot dog*, *blackmail*, *high-handed*, etc.
 - e) incorporating verbs: *eavesdrop*, *manhandle*, *boot-lick*, etc.
 - f) pseudo-idioms: *kith and kin*, *spic and span*, *to and fro*, etc. (Makkai, 1972, pp.135-172)

Makkai’s choice of ‘pepper and salt’ as an irreversible binomial is a curious one as the reverse appears to be more common according to occurrences in the British National Corpus (‘pepper and salt’ – 43, ‘salt and pepper’ – 202).

- sememic idioms:
 - a) proverbs – *don’t count your chickens before they’re hatched*
 - b) familiar quotations – *Brevity is the soul of wit*
 - c) ‘first-base’ idioms associated with a national game like baseball – *have two strikes against one*, *never to get to first base*
 - d) idioms of ‘institutionalised politeness’ such as *may I...X?* with interrogative intonation for ‘I want to...X’ or *may I ask who’s calling?* for the imperative ‘identify yourself’
 - e) idioms of ‘institutionalised greeting’ – *how do you do?*, *so long*
 - f) idioms of ‘institutionalised understatement’ – *I wasn’t too crazy about him*

- g) idioms of ‘institutionalised hyperbole’ – *he won’t even lift a finger* (Makkai, 1972, pp.172-179)

Subsequent linguists have borrowed and added to these categories. Alexander (1984; 1987) uses *ournure* idioms, irreversible binomial idioms, phrasal compound idioms, phrasal verb idioms, and adds:

- proverbial idioms: *the land of Nod*
- metaphorical/allusive idioms: *a hot potato*
- idiomatic similes: *as bold as brass*

Like Makkai, Alexander (1984, 1987, 1989) distinguishes between phrase-length (idioms: phrasal verbs, *ournures*, irreversible binomials) and sentence-length expressions (discourse-structuring devices, proverbs, catchphrases, quotations, allusions), taking into consideration their pragmatic function in communication.

Similarly, McCarthy (1992, pp.56-7; 1998, pp.130-1) also uses *ournure* idioms, phrasal verbs, binomials and trinomials, and adds:

- prepositional expressions: *in two shakes of a lamb’s tale*
- frozen similes: *(as) keen as mustard*
- possessive ‘s phrases: *a king’s ransom*
- opaque compounds: *a mish-mash*
- idiomatic speech routines, gambits and discourse markers: *by the way*
- restricted collocations: *breakneck speed*
- cultural allusions: *to be or not to be*

While some of the idiom types are obvious, based on linguistic terms, others such as Makkai’s idioms of ‘institutionalized politeness’ (*May I ask who’s calling?*) are more examples of ‘nativelike’ than idioms. In addition, while some of these types seem unambiguously identifiable (if not unambiguously idiomatic), a category such as ‘*ournure*’ “is rather like a dustbin for non-

categorizable expressions” (Strassler, 1982, p.18). New terms have been created to describe a wide variety of new phrases as they appear, but are descriptive and not supported with a strong theoretical base. One of the aims of this study is to reduce the present large number of constructions classified as idioms through a clearer and more restrictive definition, because only if the set of idioms is divided into separate and distinguishable categories does the teaching of them in an ESL/EFL classroom become a more realistic task.

2.5 CLASSIFICATIONS OF IDIOMS

As well as listing these different types of idioms, attempts have been made to classify them in a variety of ways – by semantics, syntax, or function, and these attempts will be examined next.

2.5.1 Semantic classifications of idioms

Semantic classifications have included:

- semi-idioms: phrases such as *white lie*, *swallow one’s pride*, *stir up trouble* where at least one word contributes its literal meaning (Cowie, 1981; Newmeyer, 1974; Weinreich, 1969). However, other phrases (*what on earth*, *where in God’s name*) may also have at least one word that contributes to its literal meaning, but are not necessarily idioms.
- semi-opaque idioms: where the meaning is not as easy, but may be possible, to guess: *keep one’s head above water*, *sail too close to the wind* (Fernando & Flavell, 1981). Other phrases may also have meanings that are not easy, but possible, to guess (*chances are*, *how come*) but are not necessarily considered idioms.
- pseudo idioms: where one of the elements does not make any sense on its own, as in *spic and span*, *kith and kin* (Makkai, 1969). Other phrases could also contain an

element that does not make sense on its own (*easy-peasy, blah blah blah*) but are not necessarily idioms.

- pure idioms/well-formed idioms/idioms of decoding: One claim is that ‘pure idioms’ (Cowie, 1988; Fernando & Flavell, 1981) include those that have both a literal and a non-literal meaning. These have also been described as ‘well-formed’ idioms (Chafe, 1968) and ‘idioms of decoding’ (Makkai, 1972) – *kick the bucket, spill the beans*, but are said to exclude those idioms that violate ‘truth conditions’ – *rain cats and dogs, move heaven and earth, throw caution to the winds*. This would make them ambiguous or potentially misleading (Makkai, 1972). Alternatively, Moon’s (1998a) description of ‘pure idioms’ as ‘opaque metaphors’ which cannot be understood without knowledge of the historical origins (*bite the bullet, kick the bucket, over the moon, red herring*) includes those which ‘violate truth conditions’, a contradiction of earlier definitions. In addition, other phrases can also have ambiguous meanings which can be misleading (little red book – (Mao’s) Little Red Book, (please) tell me about it – Tell me about it!, you’re not just a pretty face!) but are not necessarily idioms.
- full idioms: for which the constituent or series of constituents contains no formatives whose ordinary lexical meaning contributes to the semantic interpretation: *kick the bucket, toe the line, red herring, butter up* (Newmeyer, 1974). Similarly, other phrases may also contain no formatives whose ordinary lexical meaning contributes to the semantic interpretation (of course, by the way) but are not necessarily idioms.
- figurative idioms: a group of expressions which have figurative meanings (in terms of the whole combination in each case) but which also keep a current literal interpretation: *catch fire, close ranks* (Cowie *et al.*, 1983). Moon (1998a), however, includes ‘transparent metaphors’ (*alarm bells ring, behind someone’s back*), ‘semi-transparent metaphors’ (*on an even keel, the pecking order*) and ‘opaque metaphors’ (*red herring, shoot the breeze*) in her classification of ‘fixed expressions and idioms’. The figurative element of idioms will be dealt with later in this study, when criteria are established to define an idiom.

Table 1 Semantic classification of idioms

Yorio (1980)	Cowie, Mackin & McCaig (1983/93)	Alexander (1987)	Cacciari and Glucksberg (1991)	Howarth (1998)	Moon (1998a)	Fernando and Flavell (1981)	Fernando (1996)
Transparent (not idioms): <i>your face looks familiar</i>	Open collocation: <i>fill the sink, a broken window, in the raw</i>	Literal phrase: <i>hit the ball</i>	Analyzable transparent: <i>break the ice, spill the beans</i>	Free combination: <i>under the table</i>	Transparent metaphor: <i>alarm bells ring, behind someone's back</i>	Literal and/or transparent: <i>cut wood, break eggs, rely on, add fuel to the fire</i>	Literal idiom [sic]: <i>tall, dark and handsome, on foot, for example</i>
Semi-transparent (expressions or idioms): <i>shake hands, bumper to bumper, skyscraper</i>	Restricted collocation/ Semi-idiom: <i>jog one's memory, a blind alley, catch someone red-handed</i>	Semi-idiom: <i>hit a six</i>	Non-analyzable: <i>by and large</i>	Restricted collocation: <i>under attack</i>	Semi-transparent metaphor: <i>on an even keel, grasp the nettle, the pecking order</i>	Metaphor/Semi-transparent: <i>skate on thin ice, kill two birds with one stone, the boot/shoe is on the other foot</i>	Semi-literal idiom: <i>kith and kin, drop names</i> Semi-idiom: <i>catch your breath, foot the bill</i>
	Figurative idiom: <i>catch fire, close ranks, beat one's breast, bleed someone white</i>	Metaphorical idiom: <i>hit the jackpot</i> Figurative idiom: <i>hit list</i>	Quasi-metaphorical: <i>giving up the ship, count your chickens before they're hatched, carry coals to Newcastle</i>	Figurative idiom: <i>under the microscope</i>		Metaphor/Semi-opaque: <i>burn one's boats, tarred with the same brush, off the top of one's head</i>	
Opaque/True idioms: <i>by and large, take a leak, knock on wood, be on the wagon</i>	Pure idiom: <i>blow the gaff, kick the bucket, in a nutshell</i>	Opaque/Pure idiom: <i>hit the sack</i>	Analyzable- opaque: <i>kick the bucket</i>	Pure idiom: <i>under the weather</i>	Opaque metaphor/Pure idiom: <i>bite the bullet, over the moon, red herring, kick the bucket</i>	Full idiom/ Opaque: <i>pull someone's leg, pass the buck, trip the light fantastic</i>	Pure idiom: <i>spill the beans, chin wag, red herring, take 40 winks, have cold feet</i>

As we can see many linguists, instead of giving a narrow definition of idioms, prefer to describe a scale or continuum of idiomaticity (Bolinger, 1976; Cowie *et al.*, 1983; Fernando, 1978, 1996; Wood, 1981) from least to most idiomatic based on semantic considerations. These scales or continuums, however, including those listed in Table 1, do not appear to be based on any central understanding of what an idiom is, but rather on a perception that things are not equally transparent or opaque. The end result is some overlapping categories. While some idioms (*kick the bucket*, *red herring*) are deemed to be opaque and at the extreme end of ‘pure idioms’, others (*spic(k) and span*, *by and large*) can be listed in different categories resulting in a plethora of subtypes and a relative lack of agreement about transparency/opacity. The fact that major disagreements exist shows that, as yet, no one has reached a satisfactory situation for the language teacher or language learner trying to apply this. In other words, while there is a difference between *watch the clock* and *kick the bucket*, both of these MWUs are in the same dictionary (Cambridge *International Dictionary of Idioms*, McCarthy & Walter, 1998). Furthermore, the classifications have not helped to clarify exactly what an idiom is, nor to distinguish an idiom from a non-idiom. Establishing clear criteria should enable either the inclusion or exclusion of these various classifications from the idiom category.

2.5.2 Syntactic classifications of idioms

The syntactic classifications have included different grammatical categories:

- asyntactic idioms: idioms which are not well-formed, although some grammatical structure is normally evident – *by and large*, *far and away* (Cruse, 1986, pp.37-8)

- commonly occurring phrase patterns:
 - a) noun phrase – *a crashing bore*
 - b) adjective phrase – *free with one's money*
 - c) prepositional phrase – *in the nick of time*
 - d) adverbial phrase – *as often as not* (Cowie *et al.*, 1983, pp.xxix-xxxii)
- clause patterns:
 - a) verb + complement – *come clean*
 - b) verb + direct object – *foot the bill*
 - c) verb + direct object + complement – *bled his family dry*
 - d) verb + direct object + adjunct – *cast your net wide*

(Cowie *et al.*, 1983, pp.xxix-xxxii)
- clause/phrase patterns:
 - a) possessive clause – *got a taste of their own medicine*
 - b) noun phrase – *the common touch*
 - c) adjective phrase – *not fit to wash his feet*
 - d) prepositional phrase – *under your own steam*
 - e) adverbial phrase – *none too soon*
 - f) noun + noun pattern – *fair and square*

(Cowie *et al.*, 1983, pp.xxxiii-xxxviii)
- different grammatical types:
 - a) subject + predicator + object – *bury the hatchet*
 - b) subject + predicator + object + adjunct – *keep tabs on someone*
 - c) subject + predicator + adjunct – *come to grief*

(Moon, 1998a, pp.85-92)

These three categories accounted for almost 70% of the twelve commonest patterns found in Moon's (1998a) corpus, with the remaining nine categories accounting for a further 30%.

Various syntactic classifications have been given (Yorio, 1980; Cowie *et al.*, 1983, 1993; Fernando, 1996; Moon, 1998a) which, apart from a few imparsable structures such as *by and large*, *far and away* (Cruse, 1986, pp.37-

38), seem to include virtually any type of immediate constituent. The trouble with such syntactic classifications is that they provide descriptions of internal structure (verb + noun: *blow the gaff*; preposition + noun: *under the weather*, Howarth, 1998, p.28) rather than ‘definitions’ of idioms. While providing information on syntax can help learners see the various grammatical patterns of an idiom (Cowie *et al.*, 1983, 1993, pp.xxvii-xxxvii), the large number of patterns tends to overwhelm the learners. Establishing clear criteria to define an idiom should clarify the inclusion or exclusion of these categories in the idiom category.

2.5.3 Functional classifications of idioms

A third classification identified is that of ‘functional idioms’ (Cowie *et al.*, 1983, p.xi):

- catchphrase (often humorous) – *Did he fall or was he pushed?*
- saying (usually a comment on something) – *out of sight, out of mind*

One of the first people to do a detailed examination of one particular function of idioms – interpersonal relations – was Strassler (1982, p.116) who argued that “...when using an idiom the speaker conveys more information than its semantic content. He either establishes a social hierarchy or he tests the hearer’s opinion in this matter.” Strassler (1982) discovered that idioms deal with personal reference (first person: I can’t dance *worth shit*), reference to a communicative partner (second person: I mean, why have you got such a, a sort of a *chip on your shoulder* about it), and most commonly, to a third person (O’Brien is a *ball player*) or object (...you try to brush it off the table as being *pie in the sky*) (examples from Strassler’s, 1982 corpus). In addition, he noted the use of second person idiom for partners of higher status, and first person

idiom for partners of lower status. While it is true that ‘have a chip on your shoulder’ is seldom used with first person, it can be used with second and third person (you seem/he seems to *have a real chip on your/his shoulder* about not getting the job). Moreover, as Strassler’s (1982) study used a limited corpus of trial transcripts, therapeutic session recordings and White House transcripts, his theory would need to be tested on a more representative corpus.

Subsequent studies regarding this particular function have looked at how idioms are used in spoken English – for example, making complaints (Drew & Holt, 1988), or for topic transition (Drew & Holt, 1995, 1998; McCarthy, 1998). McCarthy (1998, p.145) notes other ways in which idioms function in discourse, including:

- Idioms are never just neutral alternatives to literal, transparent semantically equivalent expressions.
- Idioms always *comment* on the world in some way, rather than simply *describe* it. They are evaluative and frequently involve threats to face.
- Speakers are aware of the face-threatening potential of idioms and are careful to use idioms generally only for third person reference, or to mitigate threats to face in first- and second-person reference.
- Idioms are communal tokens that enable speakers to express cultural and social solidarity; this is particularly so of those that are direct cultural allusions such as proverbs, quotations, etc.
- Idioms may be relatively infrequent in occurrence, but it is possible to predict where they are likely to occur (e.g. in particular genres and at particular places).
- Idioms occur in a wide variety of forms, not just the *verb + complement* type. In our data samples we have seen idiomatic nominal compounds, frozen similes, binomials, phrasal verbs, and clausal idioms.

Moon (1998a) also looks at this, devoting an entire chapter in her book (*Fixed Expressions and Idioms in English, A corpus-based approach*) to what she terms the ‘discoursal functions’ of her ‘fixed expressions and idioms’ (FEIs). She divides these into various categories, with 79% of the ones found in her corpus from the first two categories, and 47% of them classified as having two or more text functions (Moon, 1998a, pp.215-43):

- a) informational (stating proposition, conveying information): *rub shoulders with, in the running, catch sight of something*

- b) evaluative (conveying speaker's evaluation and attitude): *kid's stuff, a different/fine kettle of fish, near the knuckle*
- c) situational (relating to extralinguistic context, responding to situation): *excuse me!, long time no see, knock it off!, talk of the devil*
- d) modalizing (conveying truth values, advice, requests, etc): *I kid you not, you know what I mean, to all intents and purposes*
- e) organizational (organizing text, signalling discourse structure): *by the way, for instance, talking of ____, be that as it may*

Simpson & Mendis (2003, pp.427-431) also found examples of this 'cross-functioning', with idioms having a variety of pragmatic functions:

- a) evaluation (much more likely to occur in a third-person context, especially if it represents a threat to face) ...the paintings didn't seem to use the conventional skills of perspective the perspective is all *out of whack* here...
- b) evaluation (under certain circumstances, face-threatening idioms can be used in a second-person context) ...you were just mean to me again, oh you've got this *evil side that rears its ugly head*...
- c) description (a function that overlaps with evaluation, an example of 'cross-functioning') the Roman empire was also acquired for economic reasons, this is very different, from European empires, uh in the eighteenth and nineteenth centuries, uh where both national pride and, uh notions of economic expansion *go hand in hand*...
- d) paraphrase (another function closely related to description and particularly well suited to academic spoken discourse given its heavy use of explication) ...so women knew that uh if they were labelled noncompliant if they were you know *put up a stink* about where they delivered their children, this might have a negative impact on health care for their whole family...
- e) emphasis (idioms can be used to emphasise content or reinforce an explanation, two functions that complement the goals of academic discourse) ...mostly in, doing data analysis we're interested in posing particular questions that're interesting to us, and we're not interested in explaining the total variance in our outcome by *throwing everything in but the kitchen sink* so, by looking only at R-Square of how good is our model that would be the kind of what I would call the *kitchen-sink* model...
- f) collaboration (idioms can be used to create collaborative discourse and establish a sense of solidarity within a group of speakers)...are they more likely to take steps if the press *puts some heat on them?* ...the press has already *put heat on 'em* they're *putting heat on themselves*...
- g) metalanguage (idioms can be used to show that the speaker is about to move in a different though not entirely irrelevant direction) ...and the Basques by the way have been some in the history of Spain, if I may just uh, *go off on a tangent* the Basques have been in Spanish history and Spanish religious history...

As both spoken and written genres have been considered, it appears that the functions of idioms vary with the genre.

2.6 A SCALE OF IDIOMATICITY

While others examined the transformational potential of idioms (Katz & Postal, 1963; Weinreich, 1969), Fraser (1970) was the first to look at it on a 'scale'

with his ‘frozenness hierarchy’. His scale involves seven levels, ranging from, at one end, L6 – Unrestricted (so no idioms belong on this level) and L5 – Reconstitution, which allow most flexibility (*pass the buck, spill the beans*), to L1 – Adjunction (*kick the bucket, shoot the bull*) and L0 – Completely frozen (*build castles in the air, trip the light fantastic*), which allow the least flexibility, at the other end. Fraser (1970, pp.39-40) argues that “the most frozen idioms, those characterized by L0, permit no distortion, those least frozen, L5, permit considerable alteration” and that “any idiom marked as belonging to one level is automatically marked as belonging to any lower level”. Examples from the BNC show that some of Fraser’s ‘completely frozen’ idioms are in fact not completely frozen (‘other things associated with *the tripping of the light-fantastic*’, *dancing among trippers on the light fantastic floor*). While subsequent linguists have developed idiom dictionaries based on a variety of classifications (Cowie et al., 1983), Fraser’s (1970) scale has not proved reliable or comprehensive enough for others to adopt.

Some authors have proposed a ‘scale of idiomaticity’ (Alexander, 1987; Bolinger, 1976; Cowie *et al.*, 1983; Fernando, 1996, Howarth, 1998) or a continuum (Cowie *et al.*, 1983; Fernando, 1978; Moon, 1998b; Wood, 1981) from least to most idiomatic based on semantic considerations. For example, Alexander’s (1987, p.110) scale or cline of expressions ranged from:

- literal phrase (*hit the ball*)
- semi-idiom (*hit a six*)
- metaphorical idiom (*hit the jackpot*)
- figurative idiom (*hit list*)
- opaque/pure idiom (*hit the sack*)

Howarth (1998, pp.27-8) used a different scale with expressions ranging from:

- free combination (*under the table*)
- restricted collocation (*under attack*)
- figurative idiom (*under the microscope*)
- pure idiom (*under the weather*)

Yet another scale focuses on transparency, from “totally transparent to totally opaque” (Fernando & Flavell, 1981, p.28), or “fully transparent” and “to some degree opaque” (Cruse, 1986, p.39). Examples of this are in Table 2 below (Fernando & Flavell, 1981):

Table 2 Fernando and Flavell's examples of transparency and opacity

Literal		Metaphors	Idioms
Transparent	Semi-Transparent	Semi-Opaque	Opaque
<i>cut wood</i>	<i>skate on thin ice</i>	<i>burn one's boats</i>	<i>pull someone's leg</i>
<i>break eggs</i>	<i>kill two birds with one stone</i>	<i>tarred with the same brush</i>	<i>pass the buck</i>
<i>bring in</i>	<i>add fuel to the fire</i>	<i>the boot/shoe is on the other foot</i>	<i>trip the light fantastic</i>

The judging of which idioms belong in which category appears somewhat subjective, as the difference in ‘transparency’ between *kill two birds with one stone* and *the boot/shoe is on the other foot* must be a matter of opinion. Lewis (1993, p.98) argues that “the transparency of idiomatic expressions is a matter of degree” making a “spectrum of idiomaticity”. Others argue that the ‘scale’ or ‘continuum’ models are merely descriptive (Wray & Perkins, 2000). Table 1 includes these categories and gives some indication of how difficult all of this is for the language teacher to explain clearly, and how confusing all of this must seem to an ESL/EFL learner.

2.7 THE LACK OF A CLEAR DEFINITION

Problems in classifying idiomatic MWUs lead to problems in presenting them usefully for ESL/EFL learners. Table 1 shows that while there may be agreement about the idiomaticity of *kick the bucket*, *red herring* and a few

other well-rehearsed examples, when it comes to the less extreme examples such as *spill the beans* or *by and large*, there is not only less agreement on the degree of transparency/opacity, but the terminology also becomes variable and not particularly helpful (*spic(k)* and *span* is variously termed irreversible binomial idiom, compound idiom or pseudo idiom). Cowie (1998) notes the same problem, commenting that differences between word combinations such as free phrases, restricted collocations and idioms – all crucial to the foreign learner – are neither presented consistently nor explained adequately in reference works. With this lack of clear agreement on exactly how to define an idiom, the problem is exacerbated rather than alleviated, and idioms become a ragbag of teaching problems with teachers and learners trying to cope with the very large collections now found in idiom dictionaries (the *Cambridge International Dictionary of Idioms*, McCarthy & Walter, 1998, for instance claims to contain ‘around 7,000 idioms’), some of which it could be argued are not idioms (*ad nauseam*, *You can say that again*, *a train of thought*, *Sarcasm is the lowest form of wit*). So learners can consult one of the various idiom dictionaries for an explanation of the MWU, but will get no real understanding of why it belongs in this specialised dictionary or what it has in common with others in the same source. Furthermore, the large number found in each idiom dictionary is overwhelming to the ESL/EFL learner who may think it is necessary or desirable to learn them all.

It has been suggested (McCarthy, 1998, p.130) that not having a clear definition can be both a positive as well as a negative thing:

But a definition of idioms that has blurred boundaries has advantages as well as disadvantages: it usefully enables us to incorporate within the term a wide range of

fixed expressions over and above clause-idioms...and idiomatic phrasal verbs...,
which are traditionally most often focused upon in language teaching materials.

However, different coping strategies may be needed for different types of MWU. As Weinreich (1969, p.26) notes:

I think we have seen ample evidence in the past few years of how much more instructive it can be to study small problems under powerful magnification than to try to take in great heaps of phenomena in sweeping surveys.

Therefore, the emphasis will be on establishing clear criteria to define an idiom and applying these criteria to the large variety of MWUs presently called idioms, with the aim of producing a smaller, more manageable number for the ESL/EFL learner to begin to understand. The previous criteria that have been used to define an idiom will now be examined in more detail.

2.8 IDENTIFYING CRITERIA FOR DEFINING IDIOMS

There are certain recurrent themes in all the definitions of idioms: compositionality, institutionalisation, and degree of frozenness/fixedness.

2.8.1 Non-compositionality:

Compositionality relates to meaning. The meaning of a construction is compositional if it is derived transparently from the meanings of its elements. Idioms are said to be non-compositional because their meanings are not the sum of the meanings of their parts (Alexander, 1978; Benson, 1985; Bolinger, 1975; Cacciari & Tabossi, 1988; Coulmas, 1981; Cowie et al., 1983; Cruse, 1986; Fernando & Flavell, 1981; Fraser, 1970; Katz & Postal, 1963; Kovecses & Szabo, 1996; Makkai, 1975; Nattinger & DeCarrico, 1992; Reagan, 1987; Schweigert & Moates, 1988; Weinreich, 1969; Wood, 1981). However, it

could be argued that other MWUs are also non-compositional (of course, in effect) and yet are not generally considered idioms. A more detailed discussion of compositionality follows in 2.9.

2.8.2 Part compositionality:

Idioms can be placed along a continuum from transparent to opaque, with idioms which have one constituent that contributes to the meaning being called ‘partly-compositional’ (Fillmore, Kay, & O’Connor, 1988; Gazdar, Klein, Pullam, & Sag, 1985; Gibbs, 1994; Gibbs & Nayak, 1989; Gibbs, Nayak, Bolton, & Keppel, 1989; Gibbs, Nayak, & Cutting, 1989; Gibbs, 1990; Lakoff, 1987; Langacker, 1986; Newmeyer, 1972; Nunberg, 1978; Nunberg, Sag, & Wasow, 1994; Ruhl, 1975; Titone & Connine, 1999; Wood, 1981). However, the transparency or opacity of such expressions is somewhat subjective, and other MWUs (so much for ____, at any rate, so to speak) may arguably be placed at different points along the transparency-opacity continuum.

2.8.3 Institutionalisation:

Being ‘institutionalised’ means being commonly used by a large number of people in a speech community (Fernando & Flavell, 1981; Michiels, 1977; Moon, 1998a). However, other MWUs are also ‘institutionalised’ or ‘lexicalised’ (come to think of it, for the time being) but are not generally considered idioms.

2.8.4 Frozenness:

How ‘frozen’ or ‘fixed’ the words that make up an idiom are varies. Katz & Postal (1963) were some of the first to clearly identify the transformational

deficiencies of idioms, later developed by Weinreich (1969) and Fraser (1970).

Descriptions have included:

- a) being frozen and idiomatic, frozen and not idiomatic, not frozen and idiomatic, not frozen and not idiomatic (Mel'cuk, 1960)
- b) exhibiting certain transformational deficiencies (Chafe, 1968)
- c) being immutable in the sense that its parts are unproductive in relation to the whole in terms of the normal operational processes of substitution, transposition, expansion, etc. (Mitchell, 1971)
- d) being partially or totally resistant to the syntactic manipulations (Michiels, 1977)
- e) transformationally deficient in one way or another (Fernando & Flavell, 1981)
- f) being composite units of differing degrees of variability (Cowie, 1981)
- g) allowing internal lexical variation, syntactic variation, and variation in word order (Sinclair, 1987)
- h) being relatively frozen expressions (Bahns, 1993)
- i) exhibiting a syntactic form which is to a greater or lesser degree fixed (McCarthy, 1992)
- j) exhibiting lexicogrammatical fixedness or formal rigidity, which implies some degree of lexicogrammatical defectiveness (Moon, 1998a)

One linguist identifies this criterion of being 'completely fixed' as essential in the definition of an idiom:

- k) being wholly non-productive in form (Wood, 1981)

As can be seen, linguists vary in regard to this criterion, from not allowing any transformation of idioms (Wood, 1981) to allowing "different *degrees* of possible fixity or 'frozenness', both syntactic and semantic" (Carter, 1987, p.58). Howarth (1996, pp.8-9) notes that many Russian phraseologists have used relative fixedness as a primary distinguishing feature but argues that the fixedness of a word combination varies from the immutable 'irreversible binomial' (*bag and baggage*) to others (greetings, gambits, formulae) which

allow some substitution of elements. Other linguists allow some variation in form, depending on collocational restriction, syntactic structure and semantic opacity (Carter, 1987; Glaser, 1988). Moon (1998a, p.120) found that rather than idioms being ‘completely fixed/wholly non-productive’ in her data of ‘fixed expressions and idioms’ “around 40% of database FEIs have lexical variations or strongly institutionalized transformations, and around 14% have two or more variations on their canonical forms”. Therefore, even idioms that may be considered by some to be ‘fixed’ are in reality used in a variety of forms:

- *down the chute / drain / pan / plug hole / toilet / tubes / tube*
- *fan the fire of something / fan the fires of something / fan the flames of something*
- *add fuel to the fire / add fuel to the flame / add fuel to the flames / fuel the fire / fuel the fires / fuel the flame / fuel the flames* (Moon, 1998a, pp.161-7)

Similarly, other MWUs also include both those that are ‘to a greater or lesser degree fixed’ (to make a [rather/very/extremely] long story short) and those that are completely fixed (for instance, all in all, on time) but are not generally considered idioms. Therefore, this criterion alone of being partly or completely fixed does not help to define an idiom, nor to distinguish an idiom from a non-idiom.

To clarify the idea of the ‘frozenness/fixedness’ of an MWU, this example will illustrate how easily ordinary grammatical constructions can be manipulated (1), as compared with an idiomatic MWU (2):

1) She saw the cat.

She saw it.

The cat, she saw.

The cat was seen by her.

She saw the pussy.

She saw the tabby pussy.

She is seeing the cat right now.

She saw cats.

2) She kicked the bucket.

She kicked it.

The bucket, she kicked.

The bucket was kicked by her.

She kicked the pail.

She kicked the plastic pail.

She is kicking the bucket right now.

She kicked buckets.

Similar attempts to manipulate idioms as in (2) cause the idiomatic meaning of the phrase to be lost, with only the literal meaning retained. It seems clear that what is central to this is the notion of compositionality. Frozenness/fixedness, compositionality, and institutionalisation are related. Something which is non-compositional cannot be produced simply as a sequence of elements (words) but must be held in memory. In other words, it becomes institutionalised and is held in memory to survive. Only compositional constructions can be created on the spot. Non-compositionality depends on prior institutionalisation (commitment to memory) and this involves some degree of fixedness. Thus, compositionality is the key criterion here.

To begin with, therefore, I accept the ‘non-compositional’ criterion while leaving aside the others. The ‘non-compositional’ criterion, however, can also be applied to non-idioms and therefore requires further explanation. The issue of ‘compositionality’ will now be examined more closely.

2.9 THE ‘COMPOSITIONALITY’ OF LANGUAGE

The principle of compositionality is commonly described as Frege’s principle, although Janssen (1994) notes that while it originated with Frege, it appears nowhere in his published work so should more accurately be called ‘Fregean’. Lyons (1995, p.204) explains the theory of compositionality:

The meaning of a composite expression is a function of the meanings of its component expressions.

For decades linguists have defined idioms as being ‘non-compositional’, explaining that the meaning of the idiom is not the sum of the meanings of its parts. The prototypical example of an ‘idiom’ is that of *kick the bucket* whose meaning of ‘die’ cannot be derived from the meaning of the elements and the construction is therefore non-compositional. However, not all non-compositional set expressions are idioms. As explained (see 2.8.1), ‘fixed expressions’ such as ‘of course’ are non-compositional in that their meanings are not the sum of their constituent parts, but they are not usually defined as idioms.

Because there remain many unanswered questions, it is important to clarify what ‘compositional’ means.

2.9.1 Problems with ‘compositionality’

Before using ‘non-compositional’ to refer to idioms, a clear understanding of what ‘compositional’ means may be necessary. Having excluded constructions such as ‘compound words’ from this study, I will use them here to make a different point, namely that compositionality is not well-defined. For example, how does it apply to a compound word like ‘birdhouse’? Does the “function of the meanings of its parts and the syntactic rule by which they combine” (Partee, ter Meulen, & Wall, 1990) give the definition of ‘a house for birds’ or ‘a house shaped like a bird’? Furthermore, does ‘Wellington bus’ mean a bus on its way to Wellington, away from Wellington but originating there, or a bus going around the city of Wellington? While it could be argued that ‘birdhouse’ is a house having a connection with birds, and ‘Wellington bus’ is a bus having a connection with Wellington, the same argument applies less well to the compound noun ‘girlfriend’ where the word could mean both ‘female friend’ or ‘female involving a romantic/sexual relationship with’, or both.

We must now consider the general postulates of Construction Grammar. One such postulate seems to be that constructions not only inherit meanings from their constituents, they may also have their own semantics and pragmatics (see e.g. Kay, 1993). If a construction provides some of the meaning of a stretch of language, we should not expect the whole of the meaning to be deducible from the parts of the construction. Yet compositionality apparently makes this assumption. So just what can be compositional is not clear. Kay (1993, p.2) discusses this in the following terms:

[S]peakers of English have to know what *red* means and that it is an adjective, and they have to know what *ball* means and that it is a noun. They have to know that adjectives can co-occur with nouns in a modification structure (as in a phrase like *red ball*), and they have to know the proper strategies for giving a semantic interpretation to such adjective-noun combinations. But they do not have to know separately, or to be told, what the phrase *red ball* means. That is something that what they already know enables them to figure out.

The reference to “proper strategies for giving a semantic interpretation to such adjective-noun combinations” requires knowledge of the semantics of the construction itself. The over-inclusiveness of non-compositionality can be illustrated with the two following constructed dialogues:

a) Husband: I’ll come home early so we can all go to the zoo.

Wife: *Oh, that’ll be nice for the children!*

b) Husband: I’m sorry, but I need the car today so you’re all going to have to walk home from the Christmas party, gifts and all.

Wife: *Oh, that’ll be nice for the children!*

Clearly, the above examples show a literal reading of the final sentence in the first dialogue and a sarcastic or ironic reading of the same sentence in the second dialogue. While the principle of compositionality may apply to literal expressions, it does not deal with sarcasm or irony. It could be argued that virtually any expression can be said sarcastically or ironically and that context is needed to distinguish the literal from the sarcastic/ironic meaning, but expressions such as these (Moon, 1998a, p.200) are seldom used literally:

a fine/pretty kettle of fish	a bright spark	happily ever after
the holy of holies	one's heart bleeds	big deal
take the cake/biscuit	pearls of wisdom	ray of sunshine
God's gift to ____	whiter than white	Tell me about it!
cry all the way to the bank	need something like a hole in the head	

To conclude, as the 'compositionality principle' does not deal well with sarcasm, irony or ambiguity, it will be set aside in favour of a 'naïve approach'.

2.9.2 The 'naïve approach' to compositionality

If we adopt a naïve approach to understanding compositionality, we must go back to basics and apply a sort of 'building block' approach, adding together the meanings of each word to see if this results in the total meaning of the MWU. If the meaning of the construction can be gained from the meaning of its elements, the meaning of the construction should remain unchanged if each of those elements is replaced by its own definition. If replacing each element in a construction with its own definition results in a change of meaning, the construction cannot have been compositional. Definitions can be taken from any reputable non-historical dictionary. To see how this works, consider the following examples:

Modifier + noun combination:

- a. red + paint = **red paint**

red: 1. any of a group of colours, such as that of a ripe tomato or fresh blood, that lie at one end of the visible spectrum, next to orange, and are perceived by the eye when light in the approximate wavelength range 740-620 nonometres falls on the retina. Red is the complementary colour of cyan and forms a set of primary colours with blue and green. (*Collins English Dictionary*, Sinclair, 1994, p.1297)

paint: 1. a substance used for decorating or protecting a surface, esp. a mixture consisting of a solid pigment suspended in a liquid, that when applied to a surface dries to form a hard coating (*Collins English Dictionary*, Sinclair, 1994, p.1121)

These two definitions, in combination, tell us what ‘red paint’ means. Now let us look at the MWU ‘red herring’ from a compositionality point of view:

b. **red:** as above

herring: any marine soft-finned teleost fish of the family Clupeidae, esp. *Clupea harengus*, an important food fish of northern seas, having an elongated body covered, except in the head region, with large fragile silvery scales. (*Collins English Dictionary*, Sinclair, 1994, p.728)

‘Red + paint’ does not yield problems within “the proper strategies for giving a semantic interpretation to such adjective-noun combinations” (Kay, 1993). Even ‘red + apple’ (where the apple is red only on the outside) is straightforward, since the rules of the construction do not necessarily imply that red is the only colour involved, but that it should be perceptually dominant. However, in the second example, ‘red’ + ‘herring’ does not equal a soft-finned fish in the red colour range. A native English-speaker knows that herrings (fish) are usually a greyish colour (unless smoked), and that the term *red herring* is an idiom unrelated to either the colour ‘red’, or the fish ‘herring’ (except in its historical explanation of a smoked fish being dragged over a trail to confuse trackers, see Goatly, 1997, p.32). As this study is based on a here-and-now ‘synchronic’ rather than a ‘diachronic’ approach to MWUs (Lyons, 1977), this knowledge will not be considered relevant. *Red herring* is thus proved by the test to be non-compositional. As shown previously, however, non-idiom phrases can also be non-compositional, so this alone is not enough to determine an idiom. Thus, the test should be applied in a particular context.

To summarise, a test for compositionality is whether replacing each word in a MWU with its dictionary definition gives the same meaning as the phrase in context. If it does, the MWU is compositional. If it does not, the MWU is non-compositional.

2.10 ESTABLISHING NEW CRITERIA FOR DEFINING IDIOMS

With regard to MWUs, I accept the following:

- a MWU can be compositional or non-compositional
- an idiom is non-compositional

However, it is not just literal language which is compositional, or idioms which are non-compositional. It has been noted that irony and sarcasm may not be compositional, and nor may be some figures of speech such as metaphor as this constructed dialogue shows:

A. What's Jack's problem?

B. He's *hot under the collar*.

Statement B could refer both to the increased body heat caused by restricted clothing, possibly leading to a rash (hence, 'problem'), making it compositional, or to the feeling of anger caused by an unexplained factor, making it non-compositional. The crucial question is whether figures of speech like 'hot under the collar' are idiomatic, or whether we need to distinguish the two. It seems that a distinction is required both theoretically and from the teaching point of view because figures of speech can be interpreted according to general cognitive principles, while idioms have to be learnt. Since proving that some piece of language is non-compositional may only show that it is figurative, we have first to rule out those expressions which are non-compositional because they are figurative.

2.11 FIGURATIVE LANGUAGE

The term ‘figurative language’, most often used in contrast to ‘literal language’, includes a number of ‘figures of speech’, in particular metaphor, simile, and irony, as well as the less common litotes, zeugma/syllepsis, synecdoche, metonymy, oxymoron, hyperbole (Ortony, 1984).

2.11.1 Defining ‘figurative’

One of the most common ‘figures’ in figurative language is the metaphor – to such an extent that the words ‘figurative’ and ‘metaphorical’ are often used interchangeably. Of course, not all figures are non-compositional: similes (It was *as dark as pitch* outside) may be perfectly compositional as may zeugma. But I am not concerned with such cases here. Because metaphor is one of the most common figures, we must take a closer look at metaphors. The two levels of metaphor, conceptual and linguistic, are outlined as follows (Deignan, Gabrys, & Solska, 1997, p.352): “The term ‘conceptual metaphor’ is used to refer to a connection between two semantic areas at the level of thought (ANGER IS HEAT)...Linguistic metaphors are the spoken or written realizations of a conceptual metaphor (I grew hot under the collar / She’s got a fiery temper).” I am going to set aside ‘conceptual metaphors’ for now as they are not relevant to my argument, and focus on ‘linguistic metaphors’. Despite the work done on metaphors in the last two decades, as Cameron & Low (1999, p.77) point out, little of that has reached applied linguistics. They outline two key approaches to metaphor, noting that non-specialists would agree that “metaphor compares two things which are actually different; that a metaphor looks like an untruth but isn’t; that making sense of metaphor can help one see things in new and often interesting ways” while more technical definitions,

explain Cameron & Low (1999, p.77) include “the bringing together of two very different concepts in a linguistic expression that encourages some meaningful transfer of sense in interpretation”. To deal with figurative language generally, my argument focuses on the element of ‘untruth’, and the notions of both linguistic and pragmatic competence. I believe that an interpretation of metaphor and other figurative language is a pragmatic reinterpretation of ‘untruth’ in known circumstances (see also ‘metaphorical transfer of meaning’, Goatly, 1997, p.96). So when you encounter something which is compositionally an ‘untruth’, it is your linguistic competence which tells you that what you hear or read is actually an ‘untruth’. But it is your pragmatic competence which allows you to interpret what you have heard or read. In the non-compositional reading of Jack being ‘hot under the collar’, the compositional untruth that Jack has increased body heat only in his neck/collar area is analysed linguistically and proves to be pragmatically unlikely in context or to break Gricean principles. It then has to be reinterpreted in a way that explains the expression in its original context. While this works for metaphor, precisely the same process can be used for other figures of speech as well:

1. *The guns waited* (Seymour, 2001) (metonymy) – Our pragmatic competence tells us that inanimate objects such as ‘guns’ cannot ‘wait’, so ‘guns’ is reinterpreted to mean ‘people with guns’.
2. *I’ve told you a million times not to do that* (hyperbole) – Our pragmatic competence tells us that it is not feasible to repeat the same instruction a million times, so this is reinterpreted to mean ‘very many’ times.
3. *This land is owned by the crown* (metonymy) – Our pragmatic competence tells us that objects such as crowns do not own other objects, so ‘crown’ is reinterpreted as the person who wears the crown or the institution represented by the crown.

As with my naïve approach to compositionality, we can now attempt a similar naïve approach to figurative language. Based on Cameron & Low's (1999, p.77) description of metaphor as something that "looks like an untruth but isn't" and "a linguistic expression that encourages some meaningful transfer of sense in interpretation", I will claim that:

Understanding figurative language involves taking a compositional untruth and extracting probable truth from it by an act of pragmatic reinterpretation.

In some cases, perhaps the majority of cases given the prevalence of metaphor, this reinterpretation will be done in terms of an image or figure (or 'image schema', Gibbs, 1992). In other cases, the reinterpretation will be done in other ways as shown above. In every case, though, some pragmatic act of turning 'untruth' into 'truth' is involved. If 'truth' seems to be claiming more than is possible here, the word 'sense' can be used in its place, which would also account for what happens in non-assertive contexts. It is the possibility of reinterpretation which is the crucial point. Rather than being straightforward and easy, sometimes this reinterpretation is only done in retrospect (see, for example, 'motivated' figurative expressions, Lakoff, 1987). This is especially true for ESL/EFL learners who are likely to need help to do this.

To summarise, a test for figurativeness is whether the MWU is first recognized as an untruth and can be reinterpreted, by use of an image or other means, to deduce the intended truth. If it can (Tom's only two but he's *as good as gold*, Sally's been given *the green light* on the project), the MWU is figurative. If it cannot (How's the job? *A piece of cake.*), the MWU is non-figurative. However, many figurative expressions are referred to as being idiomatic. The difference between figurative and idiomatic expressions will now be discussed.

In describing figurative language, ‘figures of speech’ can be:

- compositional (He smiled *sheepishly*.)
- non-compositional (They’ve been *shooting the breeze*.)
- partly-compositional (He had *a long face* when the party was cancelled.)

To summarise, a test for figurativeness is whether the MWU turns ‘untruth’ (nonsense) into ‘truth’ (sense), often by using a ‘figure’ or by building a picture in your mind that gives the same meaning as the phrase in context. If it does (Ken has a memory *like an elephant* / *like a sieve*. Tom’s *a pig!*), the MWU is figurative. If it does not (His memory isn’t what it used to be), the MWU is non-figurative. I will now look at the difference between figurative and idiomatic language.

2.11.2 Figurative language vs idioms

I will try to show how metaphor in particular, and most figurative language in general, is reinterpreted to be understood by the reader or listener. Although earlier scholars have included a ‘metaphorical’ (Alexander, 1984; Moon, 1998a) or ‘figurative’ (Fernando & Flavell, 1981) category of idioms, others have referred to ‘idiomatic phrases’ as ‘figures of speech’ (Drew & Holt, 1995, 1998), or noted the ‘metaphorical mappings’ of many idioms (Nunberg *et al.*, 1994). I am drawing a distinction between these two classes, and making the claim that figurative expressions (what I am now calling ‘figuratives’) and core idioms are distinct subtypes of non-compositional MWUs.

In comparing ‘figuratives’ and idioms, I maintain that:

- figuratives are recognised as compositionally involving an untruth which can be reinterpreted pragmatically to understand the intended truth (He’s *a small fish in a big pond*) but idioms cannot (It’s *a red herring*.)

- figuratives can be undone or unpicked to work out the meaning (He's *gone off the rails* = he's no longer on the rails = he's not moving forward in a controlled, guided way) but idioms cannot (He can't *hold a candle to* my former boss.) (or at least cannot without historical knowledge, cf *red herring*).

With regard to the educational implications, different types of MWUs suggest different kinds of coping strategies and different kinds of learning, and this will be looked at briefly in Chapter 5. In other words, while figuratives require a stretching of the known meaning of the individual words, idioms may require a more etymological/historical approach. I am not saying that all figuratives will be easy for ESL/EFL learners. However, I make the assumption that turning untruth into truth by pragmatic reinterpretation – in other words, understanding figurative language – is something which is familiar in principle to speakers of all languages (see, for example, MacLennan, 1993), so that the general point does not have to be taught, though the way the general principle is exploited in any particular language may have to be. Moreover, recent research has indicated that second language (L2) teaching and learning should pay more attention to the conceptual metaphors which underlie so much of our thought and influence so much of our language (Boers, 2000b, 2001; Boers & Demecheleer, 1998; Cameron, 1999; Deignan, 1995; Deignan *et al.*, 1997; Gibbs, 1994, 1997; Kovecses, 1997; Kovecses & Szabo, 1996; Lazar, 1996; Low, 1988, 1999; MacLennan, 1993, 1994).

It is crucial that 'figuratives' can be made sense of. Idioms cannot be. They have lexicalised so far that you cannot work out their meaning. Even in cases where they arise by a freezing of figures of speech, they have lost any overt link with their compositional meaning. That being the case, and because they are institutionalised they must be associated with a fixed meaning as part of our

linguistic competence. In other terminologies, they are multiword lexemes not open to pragmatic interpretation. The best we can do with an idiom is guess its meaning as we would guess the meaning of any other lexeme. A new definition of an idiom will be discussed next.

2.12 THE NEW DEFINITION OF AN ‘IDIOM’

Based on the arguments put forward above, I leave aside the majority of criteria for (see 2.3) and descriptions of (see 2.4 and 2.5) idioms established over the past four decades for the following reasons:

1. they are too general, so apply to an extremely large number of MWUs
2. they are unable to clearly distinguish the various types of idiom from each other or an idiom from a non-idiom
3. they fail to pinpoint the crucial element of idioms as they give features common to many idioms, rather than to identify what defines an idiom

In preference, the following two criteria will be used to define idioms which, for the purposes of this research, must be MWUs (consisting of at least two words which occur elsewhere) that are:

- non-compositional
- non-figurative

Just as previous criteria have failed to distinguish an idiom from a non-idiom, so too do these two criteria. They do, however, clearly define a new, much smaller group of MWUs which will be called ‘core idioms’, and distinguish them from other MWUs previously called idioms (that I am calling ‘figuratives’), by using an Idiom Test. This test will be explained, and applied, in the next chapter.

2.13 CONCLUSION

Words frequently combine in chunks and these chunks, or collocations, can be both open/free and restricted. Collocations differ from idioms and, for the purposes of this study, both collocations and idioms will be considered a specific type of multiword unit. Due to the confusion that exists between the terms ‘idiom’ and ‘idiomaticity’, ‘idioms’ have been distinguished from ‘idiomaticity’ by replacing the latter term with ‘nativelike’. Now all references to ‘idiomatic’ will be related to idioms. The various ‘types’ of idioms, identified over the past four decades, have been examined and described as well as some new categories which have subsequently been added. While these idiom types were previously classified by lexicographers and linguists based on semantics, syntax, and function, most dictionaries have lumped them all together. In addition, a ‘scale’ or ‘continuum’ of idiomaticity was previously proposed and despite its contradictions and lack of clarity has merits as this is how language really is. However, this lack of clarity in defining idioms and in clearly distinguishing the various types has resulted in the large number currently found in idiom dictionaries. This large number of idioms presents an enormous learning load to ESL/EFL learners, making it almost impossible for them to learn to both understand, remember and use idioms correctly.

The criteria previously used for defining idioms – compositionality, institutionalisation, and degree of frozenness/fixedness – have been examined in greater detail, with ‘compositionality’ emerging as the key criterion. This has led to a closer examination of the meaning of compositionality and non-compositionality, and finally to the adoption of a ‘naïve approach’ to

compositionality in order to clearly understand what is meant by ‘non-compositional’. The result has been the emergence of figurative language as another type of non-compositional language. Further examination of figurative language led to a clearer way of defining it, which in turn led to a comparison between figuratives and idioms. The end result has been a clarification of one of the previous criteria used for defining idioms, that they are non-compositional, and the addition of a new criterion, that they are also non-figurative.

In the next chapter, we will use these new criteria for defining an idiom (non-compositional, non-figurative) to develop an ‘idiom test’, and then apply the test to idioms in idiom dictionaries to see if it fulfils the aim of producing a smaller, more manageable number of idioms for the ESL/EFL learner to deal with.

CHAPTER 3

DEVELOPING AN ‘IDIOM TEST’

3.1 INTRODUCTION

The confusion that exists regarding idioms and the many types and classifications of them has resulted in a ‘ragbag’ collection, such as the large numbers presently listed in the various idiom dictionaries which claim to have between 3,000 and 14,000 entries (see 4.2, Table 3).

In order to fulfil the aim of this research, a new definition of an idiom has been created using two criteria only. This definition will now be applied, using an ‘idiom test’, to idioms presently listed in idiom dictionaries. The hope is that the new definition will help to distinguish one group of idiomatic MWUs from another, and result in a smaller, more manageable group of idioms to be considered for teaching to ESL/EFL learners.

3.2 THE CRITERIA FOR DEFINING AN ‘IDIOM’

Based on the arguments put forward in the previous chapter, I leave aside the majority of criteria which have been used over the past four decades for defining idioms.

In preference to the previous more general criteria used to describe idioms, the following two criteria will be used to define ‘core idioms’ which, for the

purposes of this research, must be MWUs (consisting of at least two words) that are:

- non-compositional
- non-figurative

A test to apply these criteria will be examined next.

3.3 THE ‘IDIOM TEST’

In order to distinguish ‘core idioms’ from other idioms presently found in idiom dictionaries, the following test will be used for determining whether MWUs are ‘core idioms’:

1. Is the meaning of the MWU retained if you replace each lexical word in the MWU with its own definition?

Y = compositional N = non-compositional

2. Is it possible to understand the meaning of the MWU by recognising the ‘untruth’ and pragmatically reinterpreting it in a way that correctly explains the MWU?

Y = figurative N = non-figurative

3. Is there only one word in the MWU which is either not literal or non-compositional?

Y = one non-compositional element (ONCE)

N = core idiom

If we answer NO to these three questions, the MWU is a ‘core idiom’.

If we answer YES to any of these three questions, the MWU is not a ‘core idiom’ (and may be literal, a figurative, or a ONCE; we will find some borderline cases).

However, in applying this test keep in mind that if we answer YES to the first question it will be ‘literal’ and thus the second and third questions are no longer applicable. Similarly, if we answer NO to the first question but YES to the second, the third question is no longer applicable as the construction will be a ‘figurative’. And if we answer NO to the first question, NO to the second but YES to the third, the construction will have only one element that is non-idiomatic and will be a ‘ONCE’. Some ONCEs are also ‘figuratives’, however. This will be further illustrated in 3.4.

Applying the test to MWUs presently classified as idioms

The criteria will now be applied to lists of idioms from current idiom dictionaries to determine which would be included in or excluded from the ‘core idiom’ category. The reasons for exclusion could be:

- they are deemed to be ‘compositional’ and/or literal: He has sold all his *goods and chattels* with his house.
- they are deemed to be, by our definition, a ‘figurative’: He’s *hit the nail on the head*.
- they are deemed to be a ‘ONCE’ (such as when only one ‘element’ or word of the MWU is found to be a not literal, non-compositional necessary part of an idiom, with the others being ‘compositional’, ‘figurative’, or having a literal meaning): He’s ready to have a *go*. / It’s a *curly* issue. / He kept a *straight* face.

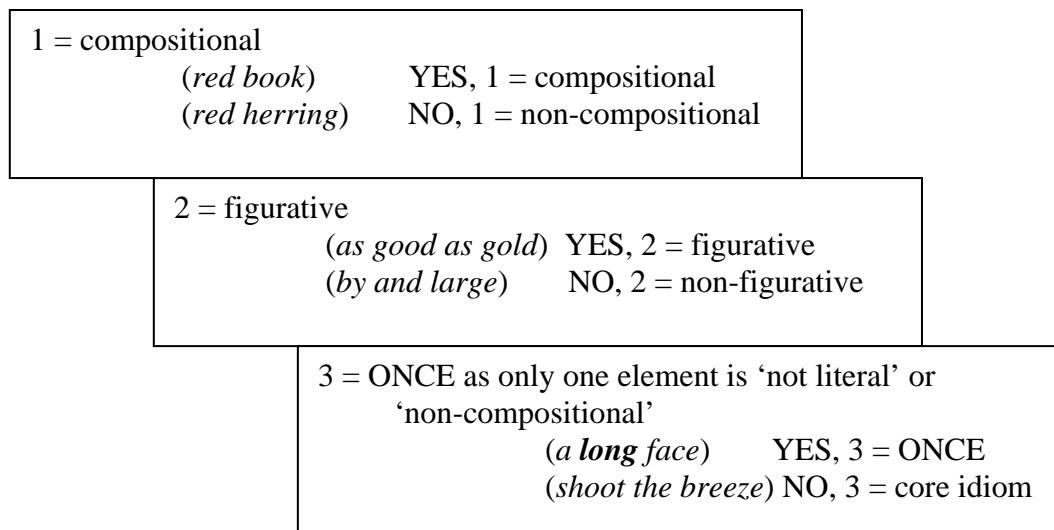
Issues to take into consideration when applying the ‘test’

A number of issues should be kept in mind:

- compound words and hyphenated words are considered single lexemes so do not meet the criteria
- foreign phrase MWUs do not meet the criteria as either the English translation would be a single lexeme, or a translation would explain the MWU
- phrasal verbs do not meet the criteria and form too large a group to deal with here so deserve separate and thorough research on their own
- the results of the test do not necessarily match intuitions: even collocational fixed phrases such as ‘of course’ would not normally be included in any idiom category, and yet meet the criteria of ‘non-compositional, non-figurative’
- how easily someone can linguistically recognise and pragmatically reinterpret the untruth of a phrase may vary according to their age, their knowledge of and exposure to these expressions, and so on. For example, people may react differently to an MWU like *a dog in a manger*, with one person seeing it as figurative, another as idiomatic. The difference would hinge on the word *manger*, the one who thought *a dog in a manger* was idiomatic associating the word purely with the Christmas story definition of *manger*, another seeing it primarily as a location for animals’ food.
- idiomatic expressions, like all units of language, are seldom encountered outside of context (unless in lists) so contextual clues may be necessary to establish what a particular untruth might be referring to (e.g., be *like a dog with two tails* – the context should be compatible with ‘great happiness/excitement’, making the connection to the dog’s tail clear).

3.4 TESTING THE ‘IDIOM TEST’

The steps will be outlined in a simple flow chart to show how the criteria were applied to each MWU.



Based on previous analysis, the hope was that the ‘fuzzy borderline’ between the two categories will be quite narrow. If too many items end up in this borderline area, it will necessitate the establishment of further criteria.

This simple table will be used to illustrate how these criteria can be applied:

Criterion 1 – YES	Criterion 2 – YES	Criterion 3 –YES
Non-relevant (literal)	N/R (figurative)	ONCE (only <i>one</i> element idiomatic)
NO – go to Criterion 2	NO – go to Criterion 3	NO – core idiom

The three criteria outlined in the ‘idiom test’ will now be applied to the following MWUs taken from page 100 of two idiom dictionaries.

3.4.1 Test 1

First, the test will be applied to the idioms listed on page 100 of the *Cambridge International Dictionary of Idioms*, McCarthy & Walter (1998):

- | | |
|---|---|
| 1. <i>in/within spitting distance</i>

<i>in/within striking distance</i> | 1 (compositional? YES, if the distance one is physically able to spit or strike is used; NO, if the distance is considered 'relatively close')

2 (able to recognise the untruth and pragmatically reinterpret it as a truth? YES)

Therefore, NOT a core idiom – literal or a figurative |
| 2. <i>drive sb to distraction</i> | 1 (compositional? YES, if defined as 'make someone distracted'; NO, if defined as 'make someone very angry or bored')

2 (able to recognise the untruth and pragmatically reinterpret it as a truth? NO, if defined as 'make someone very angry or bored')

3 (only 1 word not literal and non-compositional? YES, on the non-compositional hypothesis, <i>distraction</i> , with <i>drive</i> compositional using one of the dictionary definitions of the word)

Therefore, NOT a core idiom – a ONCE |

3. *divide and conquer/rule* 1 (compositional? YES)
Therefore, NOT a core idiom – literal
4. *pay dividends* 1 (compositional? NO)
2 (able to recognise the untruth and pragmatically reinterpret it as a truth? YES)
Therefore, NOT a core idiom – a figurative
5. *it's do or die* 1 (compositional? NO)
2 (able to recognise the untruth and pragmatically reinterpret it as a truth? YES)
Therefore, NOT a core idiom – a figurative
6. *be just what the doctor ordered* 1 (compositional? NO)
2 (able to recognise the untruth and pragmatically reinterpret it as a truth? YES)
Therefore, NOT a core idiom – a figurative
7. *dog eat dog* 1 (compositional? NO)
2 (able to recognise the untruth and pragmatically reinterpret it as a truth? YES, if understood as 'doing anything in order to succeed')
Therefore, NOT a core idiom – a figurative

8. *a dog and pony show*
- 1 (compositional? NO)
- 2 (able to recognise the untruth and pragmatically reinterpret it as a truth? YES)
- Therefore, NOT a core idiom, a figurative
9. *a dog in the manger*
- 1 (compositional? NO)
- 2 (able to recognise the untruth and pragmatically reinterpret it as a truth? YES, but see discussion above)
- Therefore, NOT a core idiom, a figurative
10. *the dog days*
- 1 (compositional? NO)
- 2 (able to recognise the untruth and pragmatically reinterpret it as a truth? NO)
- 3 (only 1 word not literal and non-compositional? YES, 'dog')
- Therefore, NOT a core idiom – a ONCE
11. *be like a dog with two tails*
- 1 (compositional? NO)
- 2 (able to recognise the untruth and pragmatically reinterpret it as a truth? YES)
- Therefore, NOT a core idiom – a figurative

12. *a dog's breakfast/dinner* 1 (compositional? NO)
- 2 (able to recognise the untruth and pragmatically reinterpret it as a truth? YES)
- Therefore, NOT a core idiom – a figurative

3.4.2 Test 2

The test will now be applied to the idioms listed on page 100 of the *Collins COBUILD Dictionary of Idioms*, Sinclair & Moon (1995):

1. *just deserts* 1 (compositional? YES/NO, depending on whether you connect 'deserts' with 'deserve')
- 2 (able to recognise the untruth and pragmatically reinterpret it as a truth? NO)
- 3 (only 1 word not literal and non-compositional? YES, on the non-compositional reading, 'deserts')
- Therefore, NOT a core idiom – a ONCE
2. *have designs on sb/sth* 1 (compositional? NO)
- 2 (able to recognise the untruth and pragmatically reinterpret it as a truth? YES)
- Therefore, NOT a core idiom – a figurative

3. *left to your own devices* 1 (compositional? NO)
 2 (able to recognise the untruth and pragmatically reinterpret it as a truth? NO)
 3 (only 1 word not literal and non-compositional? YES, 'devices')
 Therefore, NOT a core idiom – a ONCE
4. *better the devil you know* 1 (compositional? NO)
(than the devil you don't) 2 (able to recognise the untruth and pragmatically reinterpret it as a truth? YES)
 Therefore, NOT a core idiom – a figurative
5. *between the devil and the deep blue sea* 1 (compositional? NO)
 2 (able to recognise the untruth and pragmatically reinterpret it as a truth? YES)
 Therefore, NOT a core idiom – a figurative
6. *a devil of a job/the devil's own job* 1 (compositional? NO)
 2 (able to recognise the untruth and pragmatically reinterpret it as a truth? YES)
 Therefore, NOT a core idiom – a figurative

7. (*every man for himself and*) 1 (compositional? NO)
- the devil take the hindmost* 2 (able to recognise the untruth and pragmatically reinterpret it as a truth? YES)
- Therefore, NOT a core idiom – a figurative

3.4.3 Results

When applied to page 100 of the *Cambridge* idiom dictionary, none of the 12 idioms clearly passed the ‘Idioms Test’ to be a ‘core idiom’, with 9 deemed to be ‘figuratives’, 2 ‘ONCEs’, and 1 ‘literal’.

When applied to page 100 of the *Collins COBUILD* dictionary, none of the 7 clearly passed the ‘Idioms Test’ to be a ‘core idiom’, with 5 deemed to be ‘figuratives’ and 2 ‘ONCEs’.

In total, therefore, none of the 19 idioms chosen at random from the same page in two different idiom dictionaries would be considered ‘core idioms’, with the majority falling into the figuratives category.

The aim is for this ‘Idioms Test’ to be replicable, so that others may apply the same criteria. Rather than dealing with the present large collection of MWUs judged by others to be idioms, it is hoped that the test applying the new criteria will reduce the number of actual ‘core idioms’ while assigning the remaining MWUs to separate newly-created categories: ‘figuratives’ and ‘ONCEs’. The borderline may still be fuzzy because different people’s reaction to different metaphors may not be constant, as I have already illustrated. Furthermore,

some MWUs initially deemed to be ‘core idioms’ may in fact be ‘figuratives’ or ‘ONCEs’ when greater context further clarifies their meaning. The fuzziness, however, is much reduced from that which has been inherent in previous theoretical approaches, and is probably not crucial in practical terms for the language teacher. I would like to think that this is as objective as it is possible to get – but even if the tests can be improved upon, I believe progress has been made in narrowing definitions in a helpful and practical way.

3.5 CONCLUSION

The over-elaborate classifications and definitions of idioms previously given have been rejected, and the focus has been on the most important element of idioms – the fact that they are non-compositional. The aim has been to remove most of the ambiguity and give a tighter, more restrictive definition of what an idiom is. As previously noted (Cowie et al., 1983, p.xii), in regard to defining idioms: “an approach based simply on the semantic opaqueness (or transparency) of whole combinations yields a very small class of idioms. It leaves out of account, for example, an important group of expressions which have figurative meanings (in terms of the whole combinations in each case) but which also keep a current literal interpretation”. I have moved this “important group of expressions which have figurative meanings” to my newly created and much larger category of ‘figuratives’. I believe that this category, while far from problem-free for EFL/ESL learners, can be dealt with more easily by teaching the learners the conceptual metaphors which underlie our thoughts and speech and giving the learners the tools to ‘unpick’ the figurative language.

I have given an ‘idiom test’ which I hope proves to be replicable, and able to be used by others. The advantage of the test is that it fulfils the aim, resulting in a more clearly defined and therefore much reduced group of ‘core idioms’, a more manageable number for ESL/EFL learners. My aim, as explained, was to go beyond the scalar/continuum model and provide a more restrictive definition of idiom, thus allowing someone to verify whether something is or is not an idiom by getting to the ‘core’ or the heart of what an idiom is. The end result is that ESL/EFL learners should be much more aware of the group of MWUs now known as ‘core idioms’, more aware of the nature and frequency of figurative language, and better equipped for dealing with both.

With regard to having the ‘idiom test’ used by others, an inter-rater reliability check was carried out, and this will be described in more detail later (see 4.5).

The next chapter will look at steps taken to establish a definitive list of ‘core idioms’ as well as look at the categories which do not meet the core idiom criteria in order to know how best to decide which ones should be of interest to the ESL/EFL learner.

CHAPTER 4

DEVELOPING A ‘CORE IDIOM’ LIST FOR THE CORPUS SEARCH

4.1 INTRODUCTION

Large numbers of idioms can be found in the many idiom dictionaries presently on the market. However, the source of idioms in these dictionaries is unknown and if compilers of dictionaries relied on other idiom dictionaries, the possibility remains that some might have been missed. Therefore, the initial aim was to search a wide variety of sources in order to build up as comprehensive a list of core idioms as possible.

This chapter describes the procedure used to compile a comprehensive list of core idioms (including ones deemed to be ‘borderline’ – borderline/figurative [b/f] or borderline/ONCE [b/o]) before beginning a corpus search in the British National Corpus (BNC) to establish frequency. A number of steps were followed to establish a comprehensive list of core idioms, including the examination of both written and spoken sources of idioms. First it was important to clarify how the core idiom criteria would be applied.

4.1.1 Applying the core idiom criteria

To make a decision on whether the MWU qualified as a core idiom, the criteria established in Chapter 2 had to be strictly applied. The MWU had to be made up of words which occur elsewhere and be:

- non-compositional
- non-figurative

It is assumed that L1 learners applying these criteria can ‘recognise the untruth’ and ‘pragmatically reinterpret it in the known circumstances to find the intended truth’ (see 2.11.2). Therefore, even someone who has never been on a sheep farm but who understands – or has had explained – the meaning of ‘dags’ could use their pragmatic judgement to work out, in known circumstances, what *rattle your dags* means. This pragmatic judgement will be used to determine which MWUs belong in the core idiom category. Also important to keep in mind is the fact that idiomatic MWUs do not occur in isolation, but rather in context. Therefore, decisions are made on the assumption of having at least minimal context to help determine whether the MWU is a core idiom.

However, it is also worth remembering other factors, such as:

- this study is based on a here-and-now ‘synchronic’ rather than a ‘diachronic’ approach to MWUs (see 2.9.2)
- some idiomatic MWUs require a two-stage interpretation (with *The earth moved* – “something you say to describe how good a sexual experience was”, *Cambridge*, McCarthy & Walter, 1998, p.112, we can in the first stage make sense of the words, but in the second stage, we have to decide if we know how to use the MWU). Therefore, for some MWUs we may have no trouble with stage one but have some difficulty with stage two.
- MWUs such as *at loggerheads* or *to boot* include both words, so it is the full MWU that is being judged, not the single words ‘loggerheads’ or ‘boot’. Each word of the MWU has to be found elsewhere, so one could argue that ‘loggerheads’ should be excluded. However, as it occurs in a non-idiom dictionary with different meanings, it has been included as a MWU here.

Clearly, the classification of MWUs is partly subjective and may differ to a certain extent depending on who is doing the classifying. For all of these

reasons, as explained, a decision was made to ‘err on the side of caution’ and add ones which may be less well-known and less easy to put into the figuratives or ONCE categories into a ‘borderline/figurative’ or ‘borderline/ONCE’ category (see 4.4).

The first step in compiling the list of core idioms and borderlines involved examining a number of idiom dictionaries presently available on the market, and choosing one as the initial source of idioms.

4.2 CHOOSING AN IDIOM DICTIONARY

The first source of written data to be examined and the most frequently used source of idioms was that of ten idiom dictionaries which ranged in date of editions from 1986 to 2001. These ten dictionaries were further examined for information on the number of idioms, but information on exact source was not provided by any of them. The number claimed to be in each of the above ten dictionaries varied, as shown in Table 3:

Table 3 Idiom dictionaries used as a source of idioms

Name / author / date of idiom dictionary	Number of idioms claimed
<i>Cambridge International Dictionary of Idioms</i> (McCarthy & Walter, 1998)	‘around 7,000 idioms’
<i>Collins COBUILD Dictionary of Idioms</i> (Sinclair & Moon, 1995)	‘4,400 idioms’
<i>A Dictionary of Everyday Idioms</i> (Manser, 1997)	no number given
<i>English Idioms</i> (Seidl & McMordie, 1988)	‘over 3,000 idioms’
<i>Longman Idioms Dictionary</i> (Stern, 1998)	‘more than 6,000 idioms’
<i>Oxford Dictionary of English Idioms</i> (Cowie, Mackin, & McCaig, 1993, first published as <i>Oxford Dictionary of Current Idiomatic English</i> , Cowie at al., 1983)	‘over 14,000 references recorded and described’
<i>Oxford Dictionary of Idioms</i> (Speake, 1999)	‘over 5,000 idioms’
<i>Oxford Idioms Dictionary for learners of English</i> , 2001	‘more than 10,000 British and American idioms’
<i>The Penguin Dictionary of English Idioms</i> (Gulland & Hinds-Howell, 1986)	‘some 4,000 phrases’
<i>The Wordsworth Dictionary of Idioms</i> (Kirkpatrick & Schwarz, 1993)	no number given

Out of these ten dictionaries, one – the *Cambridge International Dictionary of Idioms* (McCarthy & Walter, 1998) – was selected as the main dictionary for the study, based on a number of factors:

- the number of entries claimed ('around 7,000')
- the inclusion of some source information such as 'British', 'American', and 'Australian' justifying its claim to be 'international'
- the recency of publication (1998)
- the inclusion of a frequency indicator

Having chosen the *Cambridge International Dictionary of Idioms* (hereafter referred to as *Cambridge*) as the main source of idioms, the next step involved a closer examination of it to see whether numbers were accurate and reliable before establishing an initial list of core idioms from it.

4.3 ESTABLISHING THE ACCURACY OF 'IDIOM' NUMBERS

To test the accuracy of the number claimed by *Cambridge* – 'around 7,000 idioms' – a manual count of the dictionary was carried out – first by myself and then by one other person, both of us using a conservative method of counting (one for every bold entry). The result was not the 'around 7,000 idioms' claimed but rather a total number of 5,782 separate idioms being found. It would appear that the number of idioms found in any idiom dictionary may differ from the number claimed, depending on factors such as:

- how the count was done – whether *have your finger on the pulse* and *keep your finger on the pulse* [*Cambridge*, p.135] or *go the whole hog* and *go whole hog* [*Cambridge*, p.423], *the best of both worlds* and *the worst of both worlds* [Collins COBUILD, Sinclair & Moon, 1995, p.430] were counted as single or multiple idioms

- whether the same MWU was listed more than once – *be bone dry/be dry as a bone* – under B, *be as dry as a bone* – under D [Cambridge, p.44, Cambridge, p.107]
- whether MWUs which were listed in more than one form (*a matter of sth, a matter of concern, etc., (as) a matter of course* [Oxford, Cowie et al.1993, p.381], were counted as a single or multiple idioms
- whether variations of a MWU (*keep a straight face, with a straight face, straight-faced*) [Collins COBUILD Dictionary of Idioms, Sinclair & Moon, 1995, p.132], were counted as a single or multiple idioms, and so on

The numbers claimed on the dictionary covers had therefore to be *taken with a pinch of salt*.

When an agreed upon number of idioms in *Cambridge* had been established, the criteria could then be applied to decide which ones qualified as core idioms. As explained previously (see 3.2), there are two criteria to identify a core idiom – it must be a MWU (comprised of words found elsewhere) which is both non-compositional and non-figurative. As also explained (see 3.3), while some MWUs would appear to fit comfortably in either the figurative category or the ONCE category (where only one word, or ‘element’, is non-compositional), if the MWU fits into the first category – figuratives – it is not necessary to apply the other criteria. This is one of the reasons why the figuratives category is the largest one. Therefore, of the 5,782 separate ‘idioms’ in *Cambridge*, figuratives were the largest category, followed by the ONCE category, with a considerably smaller number in the core idioms category. These three groups do not add up to the total number of idioms found in *Cambridge* as they do not include the ones that were not included because they did not meet the criteria (made up of words which occur elsewhere, non-compositional, non-figurative) to qualify as core idioms – foreign words, compound words, personal name and

fictional character MWUs, literal MWUs, and phrasal verbs, nor those judged to be borderline. We will look at these categories briefly at the end of this chapter (see 4.11). Next we will discuss why it proved necessary to have a ‘borderline’ category.

4.4 BORDERLINE CATEGORY

Any ‘judgement call’ used – often based on personal knowledge and experience – when assigning MWUs to one of the three new categories involves an element of subjectivity. For this reason, a decision was made to include a ‘borderline’ category – either ‘borderline/figurative’ or ‘borderline/ONCE’ – for those MWUs which could be judged by different people to be either core idioms, or figuratives/ONCEs.

For example, MWUs such as *butter wouldn’t melt in somebody’s mouth* or *have a bone to pick with somebody* might be judged by some who can recognise the ‘untruth’ and pragmatically reinterpret it to find the intended meanings to be figuratives, but by others who do not see any ‘figure’ to be core idioms. Moreover, it may be difficult to know whether to go *like the clappers* is to go fast, badly, or loudly, or whether to *have an axe to grind* is a good thing or not. Therefore, because the correct meaning of some MWUs used in minimal context cannot reliably and unarguably be determined by adding together the meanings of the individual words or by ‘recognising the untruth’ and ‘pragmatically reinterpreting it in known circumstances’, MWUs such as these have been put into the *Borderline* category, and judged to be either ‘borderline/figurative’ (b/f) or ‘borderline/ONCE’ (b/o). The group of MWUs

judged to be borderline is likely to vary, both in size and content, depending on how people interpret each one. While such 'borderline' idioms may simply require greater context to make the meaning clear, with the target ESL/EFL learners in mind, it seemed better to 'err on the side of caution' and add any possible borderlines to the group of core idioms to be searched for in the BNC.

However, before continuing with the search, it was important to stop and test my criteria to see whether others could successfully apply them and divide a group of idioms into the four separate categories – core idiom, figurative, ONCE, and borderline. This was done by means of an 'inter-rater reliability check', and this will be discussed next.

4.5 CONDUCTING AN 'INTER-RATER RELIABILITY CHECK'

An 'inter-rater reliability check' was carried out by two fellow ESL teachers. Both testers were given the idioms listed on identical pages (pages 80, 160, 240, 320, 400) of the *Cambridge* idiom dictionary (see Appendix 9), and asked to identify only those which qualified as core idioms. Time was spent giving each tester some minimal 'training' by explaining how the definitions applied to a variety of idioms from other pages in the dictionary, and then seeing if the testers could apply the criteria to decide which of the test group of idioms qualified as core idioms. The testers were then left with the selected pages from the idiom dictionaries and asked to return their lists after completing them. The inter-rater reliability testers separately confirmed the reliability of the criteria to define the core idioms with a high 91% agreement.

However, as the first inter-rater reliability check was conducted several months previously, a second check, with two different language teachers, was conducted in late 2001/early 2002, to once again see whether the dictionary idioms were classified accurately enough to get a similar agreement by two new testers. Once more, idioms were taken from different, but identical for each ‘tester’, pages of two separate idiom dictionaries (pages 150, 250, 350). This time, however, language teachers from outside of Auckland were selected so no ‘training’ took place, other than the written instructions (see Appendix 10). Also, this time the ‘testers’ were asked not just to identify core idioms, but also to put all the idioms into one of four categories: figuratives, ONCEs, core idioms, or borderline. The results of the second check showed an 82% agreement on which of the four categories the MWU belonged in, with lack of certainty evident in the larger number of ‘borderline’ entries. This result was not as high as the testers from the first inter-reliability test who received some ‘training’ and only had to decide one category, and actually meant that as many as one in every five MWUs was judged differently.

4.5.1 Interpreting idioms

The above result supports the claim that the interpretation of idioms is partly subjective. Some of the differences in judgement involved the ‘figuratives’ and ‘ONCE’ categories, and we have already seen that many ONCEs are also figuratives (e.g. ‘a *long* face’). However, other variables affect idiom interpretation. For example, previous research has revealed cognitive style differences, where “imagers” may be better able than “verbalisers” to associate a mental image with an idiom and thus recognise its figurativeness (Boers & Littlemore, 2000). Also, once the meaning of an idiom is known, the MWU is

likely to become a ‘figurative’ since the learner can now try to work out the meaning through an image (see Boers & Eyckmans, in press). Moreover, the etymological/historical knowledge is relevant to learners using an image to work out or remember the meaning as well, such as the two examples of *bee’s knees* and *go by the board* from the attached CD:

- ***Bee’s knees*:** This alludes to the manifestly important nectar-gathering sacs located at the point on a bee’s leg which equates to the human knee. (Donald 1994:29)
When bees climb inside the cup of a flower, pollen sticks to their bodies. The bees then carefully comb this off and transfer it to pollen sacks on their back legs. Some authorities believe that the expression refers to the delicate way bees bend their knees as they perform this operation. Rees, however, makes a strong case for an alternative theory. He argues that, although there has long been a preoccupation with bees and their knees, which has given rise to the variety of expression over the last two hundred years or so, the phrase under discussion here only dates back as far as the 1920s when it was coined as an amusing rhyme. He points to the importance of rhyme, assonance and alliteration in the origin of many expressions and a vogue in the twenties for combining features of the body or articles of clothing with parts of animals to bizarre effect. Thus we also find *the cat’s miaow*, *the cat’s pyjamas*, *the eel’s heel*, *the elephant’s instep*, and many more. (Flavell and Flavell 1992:20-1)
- ***(go) by the board*:** The *board* here is the board of a ship, and the expression has been used since about 1630 to signify something that has fallen overboard and been carried away. [Mid-1800s] (Ammer 1997:95)
The boards in this expression are those which make up the side of ship’s hull; all the rubbish from the vessel is thrown past these. (Donald 1994:150)
A ‘board’ is the side of a ship, and something which *goes by the board* falls over the side (*overboard*) into the sea and is probably lost forever. The phrase is often applied to ideas, projects or values which are discarded through impracticality or in favour of something else. (Flavell and Flavell 1992:37-8)

Other differences in interpretation may simply have to be attributed to differences in the testers’ experience and understanding of some of the idioms involved as well as the limitations of the definitions used in this study. Despite these differences, the result of the inter-rater reliability check was still considered acceptable and demonstrated that the criteria to determine the three categories could be more or less successfully applied. It also demonstrated, however, the importance of ‘hands-on’ training when preparing for an inter-reliability check, and that the subjective nature of idiom interpretation sometimes needs to be complemented by a more psycholinguistic approach.

In addition, the judging of which categories these MWUs belong in has been done so far only by English (L1) speakers. For an ESL/EFL learner, the figurativeness of an idiom is likely to depend on many more variables such as their L2 vocabulary knowledge (including polysemy), their L1 vocabulary knowledge and whether this helps or hinders their understanding of English idiomatic MWUs, and their knowledge of the L2 culture (see Charteris-Black, 2002 and Boers & Demecheleer, 2001 for evidence of the effect of cross-cultural variation on idiom interpretation). Having shown that the criteria could be applied to divide ‘idioms’ into the three new categories, it was time to establish a list of core idioms from *Cambridge*.

4.6 THE CORE IDIOM AND BORDERLINE LIST

When the criteria were strictly applied to the MWUs in *Cambridge*, a total of 58 core idioms were identified as shown here in Table 4:

Table 4 The 58 core idioms found in *Cambridge*

<i>Cambridge</i> core idioms		
<i>(be) all in</i>	<i>cut a dash</i>	<i>pull the other one / leg (it's got bells on)</i>
<i>Beat it!</i>	<i>sweet Fanny Adams</i>	<i>out (-) and (-) out</i>
<i>(be) the bee's knees</i>	<i>chew the fat</i>	<i>(be) all of a piece</i>
<i>(be) beside yourself</i>	<i>a French letter</i>	<i>(be) a piece of cake / piss</i>
<i>(a bit of) how's your father</i>	<i>cook sb's goose</i>	<i>take the piss</i>
<i>go by the board(s)</i>	<i>out (-) of (-) hand</i>	<i>come the raw prawn</i>
<i>push the boat out</i>	<i>(sb) eat your heart out!</i>	<i>(catch/have sb) bang / dead to rights</i>
<i>Bob's your uncle</i>	<i>red herring(s)</i>	<i>get your rocks off</i>
<i>make no bones about it</i>	<i>(go) hell for leather</i>	<i>cut the rug</i>
<i>shoot the breeze</i>	<i>a fine/pretty kettle of fish</i>	<i>serve sb right</i>
<i>kick the bucket</i>	<i>swing the lead</i>	<i>(be) at sixes and sevens</i>
<i>like nobody's business</i>	<i>pull sb's leg</i>	<i>to a tee</i>
<i>by and large</i>	<i>trip the light fantastic</i>	<i>take sb to task</i>
<i>not hold a candle to sb/sth</i>	<i>(be) at loggerheads</i>	<i>touch (-) and (-) go</i>
<i>(be) the cat's whiskers</i>	<i>So long!</i>	<i>cold turkey</i>
<i>pop your clogs</i>	<i>take the mickey</i>	<i>Uncle Sam</i>
<i>make sth out of whole cloth</i>	<i>not cut the mustard</i>	<i>(be) off the wall</i>
<i>come a cropper</i>	<i>(be) nip and tuck</i>	<i>and what have you</i>
<i>eat crow</i>	<i>do your nut</i>	-
<i>cut up rough</i>	<i>odds and sods</i>	-

Added to this number were the 12 MWUs put into the borderline category and judged to be borderline figuratives (b/f) or borderline ONCEs (b/o), listed here in Table 5:

Table 5 The 12 borderline figuratives/ONCEs found in *Cambridge*

<i>Cambridge</i> borderline/figuratives and borderline/ONCEs		
<i>have an axe to grind</i> (b/f)	<i>butter wouldn't melt in sb's mouth</i> (b/f)	<i>cut no ice with sb</i> (b/f)
<i>have bats in the belfry</i> (b/f)	<i>lay an egg</i> (b/f)	<i>gird [up] your loins</i> (b/f)
<i>sell sb a bill of goods</i> (b/f)	<i>a white elephant</i> (b/f)	<i>give sb the bird</i> (b/o)
<i>have a bone to pick with sb</i> (b/f)	<i>put your foot in your mouth / in it</i> (b/f)	<i>like the clappers</i> (b/o)

This took the total of core idioms and borderlines from *Cambridge* to 70. The next stage involved surveying other idiom dictionaries to see which core idioms could be added to the *Cambridge* list.

4.7 ADDING TO THE CORE IDIOMS LIST – OTHER IDIOM DICTIONARIES

A total of nine additional idiom dictionaries were used, with a thorough search carried out on seven of the remaining nine dictionaries (see 4.2), which contained numbers varying from 'over 3,000' to '14,000' (see Table 3).

- *Oxford Dictionary of English Idioms* (Cowie, Mackin & McCaig, 1993)
- *Penguin Dictionary of English Idioms* (Gulland & Hinds-Howell, 1986)
- *The Wordsworth Dictionary of Idioms* (Kirkpatrick & Schwarz, 1993)
- *Collins COBUILD Dictionary of Idioms* (Sinclair & Moon, 1995)
- *Oxford Dictionary of Idioms* (Speake, 1999)
- *Longman Idioms Dictionary* (Stern, 1998)
- *Oxford Idioms Dictionary* (no author/editor given, 2001)

The other two dictionaries were perused to see whether any core idioms had been missed:

- *A Dictionary of Everyday Idioms* (Manser, 1997)
- *English Idioms* (Seidl & McMordie, 1988)

It became clear that different idiom dictionaries listed the same idiom with different spellings, in particular hyphenation (see 4.7.2).

4.7.1 Core idioms found in other idiom dictionaries

When the other nine idiom dictionaries were surveyed, a total of 23 other core idioms were found. Many idiom dictionaries listed the same idioms, so the additional core idioms were taken from the following five dictionaries, as shown in Table 6 below:

- WW (*The Wordsworth Dictionary of Idioms*, Kirkpatrick & Schwarz, 1993)
- Ox99 (*Oxford Dictionary of Idioms*, Speake, 1999)
- Long (*Longman Idioms Dictionary*, Stern, 1998)
- Peng (*The Penguin Dictionary of English Idioms*, Gulland & Hinds-Howell, 1994)
- Ox93 (*Oxford Dictionary of English Idioms*, Cowie, Mackin & McCaig, 1993)

Table 6 Core idioms added from other idiom dictionaries

Additional core idioms found in other idiom dictionaries		
<i>The Big Apple</i> (WW)	<i>(be) off your face</i> (Ox99)	<i>take the rise out of sb</i> (Long)
<i>to boot</i> (Ox99)	<i>buy the farm</i> (Ox99)	<i>half seas over</i> (Ox93)
<i>Box and Cox</i> (WW)	<i>all gas and gaiters</i> (Ox99)	<i>hop the stick / twig</i> (Ox99)
<i>by and by</i> (WW)	<i>the ghost walks</i> (Ox99)	<i>tie one on</i> (Ox99)
<i>bust sb's chops</i> (Ox99)	<i>splice the mainbrace</i> (WW)	<i>over the transom</i> (Ox99)
<i>clean sb's clock</i> (Long)	<i>Old Glory</i> (WW)	<i>and what not</i> (Ox99)
<i>heap coals of fire on sb's head</i> (Ox99)	<i>peg it</i> (Long)	<i>wear the green willow</i> (Ox99)
<i>set sb/sth by the ears</i> (WW)	<i>part brass rags with</i> (Ox99)	-

In addition to the above core idioms, a further 5 MWUs from the same source (*Oxford*, Speake, 1999) were deemed to be borderline figuratives (b/f) or borderline ONCEs (b/o), shown here in Table 7:

Table 7 Borderline figuratives/ONCEs added from other idiom dictionaries

Additional borderline figuratives/ONCEs found in other idiom dictionaries		
<i>all ends up</i> (b/f)	<i>knock into a cocked hat</i> (b/f)	<i>go bung</i> (b/o)
<i>the bitch goddess</i> (b/f)	<i>make your number</i> (b/f)	-

Furthermore, variations were found in other idiom dictionaries and added to the *Cambridge* core idioms (with **bold** indicating zero occurrences), including:

- *come a cropper/pur**ler*** (Ox99), *chew the fat/r**ag*** (Ox99), *do your nut/bun/prunes* (not found in dictionaries but known by two others), *to a tee/T* (WW)

These variations did not add MWUs to the core idiom list as they were variations of ones already on the list. This meant a total of 28 core idioms and borderlines were added to the 70 from the *Cambridge* list, raising the total to 98. However, one thing that became evident as a result of surveying all the idiom dictionaries is the inconsistency with regard to the use of hyphens. This will be discussed next.

4.7.2 Hyphenated and non-hyphenated MWUs

Variations in the idiom dictionaries included whether the MWU was hyphenated or not. These inconsistencies with the spelling of hyphenated words led to an attempt to find information about the rules of hyphenation.

With regard to the spelling, while it was difficult to find a definite ‘rule’ concerning hyphenation, Trask (1997, p.60) sums up the generally accepted practice:

The hyphen is also used in writing compound words which, without the hyphen, would be ambiguous, hard to read or overly long. Here, more than anywhere else in the whole field of punctuation, there is room for individual taste and judgement.

Hyphens are generally used in compound modifiers if deemed necessary for comprehension (Trask, 1997), and it became clear that this “individual taste and judgement” applied to whether individual MWUs in all three categories were hyphenated or not. Therefore a decision was made to search the corpus for all hyphenated MWUs that qualified as core idioms.

Hyphenated words from other dictionaries – made up of words found elsewhere and qualifying as core idioms – were added to the *Cambridge* core idiom list to see whether they occurred in their non-hyphenated form which they did, such as Table 8 illustrates:

Table 8 Hyphenated words now included from other dictionaries

<i>to a fare-thee-well</i> (Ox99)	<i>so (-) and (-)</i> <i>so</i> (WW)	<i>cock (-) a (-) hoop</i> (Peng)	<i>such (-) and (-)</i> <i>such</i> (WW)
--------------------------------------	---	--------------------------------------	---

This increased the total number of core idioms by four. The 23 additional core idioms, 5 borderline and 4 hyphenated MWUs added another 32 to the original list of 70 core idioms and borderlines derived from *Cambridge*, raising the total number to 102. It was now time to search for other written and spoken sources of core idioms.

4.8 WIDENING THE SEARCH

It was important to come up with a systematic way of looking for other possible sources of core idioms. When trying to establish different ways of expressing temporal frequency (*once in a while, now and again, every so often*), Kennedy (1987) used a triangulation method which consisted of looking for expressions in three different places:

- a mini-corpus (*Newsweek* – every fifth page over four weeks during 1980, and *New Zealand: A New Geography* – the first 70 pages) which produced 43 different devices
- the *Concise Oxford Dictionary* which added a further 43 devices
- colleagues – ten English teachers – who were asked to think of all the ways they could to express temporal frequency, raising the total number of linguistic devices expressing temporal frequency to 291

In a similar way, I attempted to find as many ways as possible to add to the original list of core idioms from *Cambridge*. Having examined the other idiom dictionaries, the next step involved examining both written and spoken sources of idioms, including the following:

1. English language teaching course books
2. Vocabulary teaching books
3. Journal articles and book chapters written about idioms
4. Magazine horoscopes
5. Newspaper sports reports

The ‘spoken’ sources included:

6. Scripted television
7. Unscripted television

While ‘scripted television’ should more accurately be called ‘written to be spoken’, it is included in the spoken categories. First other written sources were identified and examined, followed by a similar treatment for spoken sources.

4.9 ADDITIONAL SOURCES OF CORE IDIOMS

With the target group for this study being second-language (L2) students learning English in New Zealand, the sources examined were ones considered most relevant to these learners. With this in mind, the first additional written source examined was English teaching course books.

4.9.1 English teaching course books

As a first step the English language specialist at the University Book Shop (UBS) in Auckland was consulted to find out which English course books were

the biggest sellers for language teaching in New Zealand. Five course books were identified and the levels presently on the market were examined. Table 9 lists the ones included:

Table 9 Some English teaching course books popular in Auckland

Textbook & Author(s)/Years	Levels available
<i>Matters</i> (Bell & Gower, 1991, 1992, 1995, 1999)	Advanced, Upper Intermediate, Intermediate, Pre-Intermediate
<i>New Headway</i> (Soars & Soars, 1998a, 1998b, 2000a, 2000b)	Upper Intermediate, Intermediate, Pre-Intermediate, Elementary
<i>Language in Use</i> (Doff & Jones, 1994, 1997, 1999, 2000)	Upper Intermediate, Intermediate, Pre-Intermediate, Elementary
<i>Cutting Edge</i> (Cunningham & Moor, 1998, 1999)	Upper Intermediate, Intermediate
<i>Reward</i> (Greenall, 1994, 1995, 1996)	Upper Intermediate, Intermediate, Pre-Intermediate

Not surprisingly, the majority of ‘idioms’ from these course books were figuratives or ONCEs, and were well-known ones already found in either *Cambridge* or one of the other idiom dictionaries (*have time to kill, put a brave face on, rack your brains*). Therefore, examination of these course books resulted in no new core idioms being added to the list. The next step involved examining vocabulary teaching books.

4.9.2 Vocabulary teaching books

Books aimed at teaching vocabulary to ESL/EFL learners were examined to find out which idioms were chosen by the authors/publishers to teach. Table 10 names vocabulary textbooks sold at the University Book Shop (UBS) in Auckland in 2001:

Table 10 Some English teaching vocabulary books popular in Auckland

Textbook	Author & Year
<i>Advanced Vocabulary and Idiom</i>	Thomas, 1989, 1995
<i>English Vocabulary in Use – Upper-Intermediate & Advanced</i>	McCarthy & O'Dell, 1994
<i>English Vocabulary in Use – Pre-Intermediate & Intermediate</i>	Redman, 1997
<i>English Vocabulary in Use – Elementary</i>	McCarthy & O'Dell, 1999
<i>Idioms Organiser</i>	Wright, 1999
<i>Collins COBUILD Idioms Workbook</i>	Goodale, 1995
<i>English Idioms in Use</i>	McCarthy & O'Dell, 2002

Once again, many of the idioms listed in vocabulary books were ones already found in idiom dictionaries. One book (Thomas, 1995), however, had a number of hyphenated/compound words, referred to as ‘Alliterative Expressions’, but these did not meet the core idiom criteria (see Table 18). Another book (Wright, 1999) included idioms based on a metaphoric theme, but once again the ones included – already covered in dictionaries – mostly fitted into the figuratives or ONCE categories, with no new core idioms being added to the list. As well as teaching books, numerous articles as well as book chapters have been written about idioms, and these were examined next.

4.9.3 Journal articles and book chapters written about idioms

A large number of journal articles and book chapters focusing on idioms were perused to find out which ones were identified in them. The aim was to find another source – in this case, an academic one – which might cover MWUs not already found in the other sources. A total of 128 articles or book chapters by different authors were examined (see Appendix 11), resulting in an additional 2 core idioms (*row crosshanded*, *lay out in lavender*) being added, raising the total of 102 to 104. As well as written sources which focused on idioms, another written source with a totally different focus was examined, such as horoscopes.

4.9.4 Horoscopes

Then, still from the written genre, horoscopes – previously identified as being ‘idiom’-rich (McCarthy & Carter, 1994) – were chosen. Table 11 lists these, taken from two popular weekly women’s magazines in New Zealand for the year 2000 as well as from the daily newspaper in Auckland:

Table 11 Sources and dates of horoscopes included

Source	Dates
<i>The New Zealand Women's Weekly</i>	Second week of each month, 2000
<i>Woman's Day</i>	Last week of each month, 2000
<i>The New Zealand Herald</i> ("Starwatch")	January to April, 2001

While the *Herald* also had horoscopes 'on line', these were not included because of possible repetition of the hard copy ones. All of these horoscopes were then scanned into a database and idioms were collected from these. These sources were then examined with the results that while they were full of figuratives and ONCEs (*in leaps and bounds*, *reinvent the wheel*, *your comfort zone*, *strapped for cash*), no new core idioms were added to the original *Cambridge* list. Another written source with no focus on idioms, that of newspaper sports reports, was also examined.

4.9.5 Sports reports

An unidentified source of idiomatic MWUs, that of newspaper sports reports, was chosen. Sports reports (supplied by a fellow PhD student from the Victoria University of Wellington) from roughly equal sizes of data (approximately 11,000 words each) from four New Zealand newspapers printed during 2000 were used, as shown here in Table 12:

Table 12 Newspaper sources of sport reports

Newspapers used	Location and Date
<i>The New Zealand Herald</i>	Auckland, 2000
<i>The Dominion</i>	Wellington, 2000
<i>The Wanganui Chronicle</i>	Wanganui, 2000
<i>The Otago Daily Times</i>	Dunedin, 2000

Once again, like horoscopes, these proved to be rich in both figuratives and ONCEs (*well off the pace*, *dead ready*, *short on household names*, *oozing with talent*), as well as compound words (*a hat trick*), but no new core idioms were added to the original *Cambridge* list.

All of the sources listed above are from written texts. Written language is much easier to collect, as the BNC proves with its collection of 100 million words divided into 90% written and only 10% spoken English. To widen the scope and include samples of idioms from spoken English, two additional sources were examined.

4.9.6 Scripted conversational television programmes

Scripted television programmes – more accurately a ‘written to be spoken’ source – were selected because they were thought likely to contain conversation which aimed to imitate natural conversation, a source considered ‘idiom’-rich. The scripted television programmes included soap operas, weekly and daily, as well as dramas, situation comedies (sit coms) and advertisements. Table 13 below lists the programmes selected from the months of May and October 2000:

Table 13 Scripted television programmes used for collecting idiomatic MWUs

Names of television programme	Types of television programmes
<i>Shortland Street</i>	N.Z. daily soap opera
<i>Coronation Street</i>	British bi-weekly soap opera
<i>Blue Heelers, Water Rats</i>	Australian weekly police dramas
<i>Cold Feet, At Home with the Braithwaites, Hearts and Bones</i>	British weekly sit coms
<i>Street Legal</i>	N.Z. weekly police/legal drama
<i>Montana Sunday Theatre</i>	British weekly drama
<i>Advertisements</i>	Shown during programmes being viewed

MWUs taken from these programmes were divided into the three categories. A number of different figuratives, ONCEs, idiomatic phrasal verbs or compound words, non-MWU words and slang were found, as shown:

- *best to leg it, sorted in a twinkling, the scum of the earth, brain-dead, sod it, a chat-up line, had the hots for, brownie points, you randy rabbit, cosy-up, get all touchy feely, shut your face, (she's) a ticking bomb, whiz-bang, the university of hard knocks, a tub of lard in a cardigan, Go figure!, a woman with balls*

- *bubbly, a fiver, a fag, in a tizz, skint, (my job's) shite, ratbags,(bloody) mongrel, shoot (=go ahead and speak), Bingo!, clocked [hit] him ,(little) turd, bollocks!,(little) git, bonkers, twerp, prat, a stiff [dead man], topped [killed] herself, your shout, (he's a) tosser, sannie [sandwich], daft talk, the punters, snogging, this malarky, stropky, (don't give me any of your) flannel,(got) hammered, swotting, wankers!, rugrats, a doofus, (how you get your) jollies, (having a) barney, (I'm) shattered, barking [crazy], (giving you) gyp, a nark, Greaser!, swotty, spunky, blabbed, Bugger!, twit, bottle [courage], (shut your) gob, scumbag, shonky, dork, chuffed, stoked, (some free) plonk, shoddy, the goss, a bollicking, schmoozing, grotty, Blimey!, a bimbo, Crumbs!, it's a goer/a goner*
- *mouthng off, gobbing off, psyche me out, tag along, (he's) strung out, toddle along, (I was) stitched up, muscling in, clued up, gadding about*

However, despite this usage, no new core idioms were added to the original core idioms list. The next step was to examine a similar amount of data from unscripted television.

4.9.7 Unscripted conversational television programmes

As mentioned (see 4.8), many of what I am calling 'unscripted' television programmes, such as the news, are in fact 'scripted' with the exception of comments made by newsreaders to each other, to weather or sports reporters, and so on. Despite news readers reading from television scripts, the language is written to be spoken to the wide viewing audience. In addition to the news, other unscripted television programmes included interviews, current affairs, consumer rights, and human interest programmes (cooking, gardening, and house decoration/renovation). Unscripted television programmes from the months of May and October 2000 included those in Table 14 below:

Table 14 Unscripted television programmes used for collecting idiomatic MWUs

Names of television programme	Types of television programmes
<i>News</i>	Daily N.Z. 6 pm and 10:30 pm
<i>Face the Nation, Crossfire</i>	Weekly N.Z. current affairs programmes
<i>Backchat</i>	Weekly N.Z. arts programme
<i>60 Minutes</i>	Weekly US/Australian/N.Z. current affairs
<i>Maggie's Garden Show, Ground Force</i>	Weekly N.Z. gardening/garden makeover programmes
<i>The Money Doctor</i>	Weekly N.Z. budgeting programme
<i>Changing Rooms</i>	Weekly N.Z. decorating/renovating programme
<i>Reel Life, Fat Free</i>	Short series or one-off British documentaries
<i>Country Calendar</i>	Weekly N.Z. farming/rural programme
<i>The Naked Chef</i>	Weekly British cooking show
<i>Taste N.Z.</i>	Weekly N.Z. cooking show

The same treatment was given to unscripted television as was given to scripted television programmes. As with scripted television, examples of different figuratives, ONCEs, phrasal verbs and compound words/slang were found, as shown below:

- *my emotional baggage, the cadillac of, stuff the kids, do their thing, (he's) mad as, fairy godmother, a slice of heaven, you're full of it, a bit of a dag, HP junkie, putting on your poker face, blah blah blah, put the jug on, keep on keeping on, a dash for cash, a giggle a minute, a pressure-cooker situation, the ugly duckling*
- *dork, barbie [barbeque], knackered, jammy, gob, shag, flash, grub, skint, sweet! , chuffed, hey presto, posh, bugger-all, a cock-up, old farts, bullshit, a freebie, (I'm) stoked, (I'm) rapt, hard graft, (drinking) grog, a chook, headhunted, (the rest are) toast, gizmos and gadgets, petrol-heads, the footy, a bit of R & R, (he's looking the) goods, razzamajazz, the more PC term, you rotter, a sop, yap about it, his rellies [relatives], hard yakka, the townies, (see if they're) kosher*

As well as being full of figuratives, television conversation – both scripted and unscripted – also uses a lot of sarcasm ('Thanks a bunch!', 'Oh charming!', 'a couple of rocket scientists') and ellipses ('Cool as!', 'Easy as!'). These, while not meeting the criteria for this research project, could be useful for language teaching. However, with regard to core idioms, no new ones were identified

to add to the original core idioms list. This left the total number of core idioms and borderlines at 104.

4.10 THE FINAL CORE IDIOM LIST

To summarise, as a first step idiom dictionaries were examined to choose one – the *Cambridge International Dictionary of Idioms* – as the initial source dictionary. A list of core idioms from *Cambridge* was then drawn up, with some excluded hyphenated words later being re-included, giving a total of 70 core idioms and borderlines. Next, other sources of core idioms were examined, beginning with the remaining nine idiom dictionaries which resulted in an additional 32 core idioms and borderlines being added to the list, increasing the total to 102. Following this, an exhaustive examination of first written sources of idioms (course books, vocabulary books, articles/book chapters on idioms, magazine horoscopes and newspaper sports reports) resulted in only a further 2 core idioms being added. Then to widen the scope, core idioms from the spoken – written to be spoken – sources of both scripted and unscripted television were added, but resulted in no new core idioms or borderlines being added to the list. This made an overall total of 104 core idioms and borderlines for the corpus search, shown in Table 15 below.

Table 15 Core idioms and borderlines added to original list from other sources

Orig dict.	Other dict.	Engl. teach. course books	Vocab books	Journal articles / book chapters	Mag. H'scopes	Sports reports	Script. TV	Unscript TV
70	32	0	0	2	0	0	0	0

The aim was to make this search list of core idioms as comprehensive as possible, *leaving no stone unturned* in the search of the BNC in the attempt to establish frequency of core idioms for consideration of teaching to ESL/EFL

learners. While I have attempted to compile as comprehensive a list as possible, some MWUs may have been missed. It is hoped that further research could confirm the present list of core idioms, or add any additional ones.

Having looked at the categories of idiomatic MWUs which were included in the core idioms list, it is now time to look at those categories from *Cambridge* and elsewhere which did not meet the core idiom criteria.

4.10 MWUs FROM CAMBRIDGE NOT MEETING THE CORE IDIOM CRITERIA

Certain categories of MWUs from *Cambridge* do not meet the core idiom criteria – that is, they must be comprised of words which occur elsewhere and be non-compositional and non-figurative. For example, the *Cambridge* MWUs in Table 16 contain words which do not occur elsewhere:

Table 16 Some *Cambridge* MWUs containing words not found elsewhere, so not meeting the core idioms criteria

<i>to the <u>nth</u> degree</i>	<i>have <u>dibs</u> on sth</i>	<i>fair <u>dinkum</u></i>
<i>high <u>dudgeon</u></i>	<i>come a <u>gutser</u></i>	<i>play <u>hooky</u></i>
<i>put the <u>kibosh</u> on sth</i>	<i>not on your <u>nelly</u></i>	<i>be in a <u>tizz</u></i>

Other *Cambridge* MWUs did not meet the criteria because they were comprised of foreign words, as shown in Table 17.

Table 17 Some *Cambridge* foreign word MWUs not meeting the core idioms criteria

<i>ad hoc</i>	<i>carte blanche</i>	<i>a je ne sais quoi</i>
<i>ad nauseam</i>	<i>a cause celebre</i>	<i>joie de vivre</i>
<i>apres-ski</i>	<i>deja vu</i>	<i>the nouveau riche</i>
<i>a priori</i>	<i>c'est la vie</i>	<i>persona non grata</i>
<i>au fait</i>	<i>a fait accompli</i>	<i>piece de resistance</i>
<i>ad infinitum</i>	<i>a faux pas</i>	<i>a quid pro quo</i>
<i>bona fide</i>	<i>in situ</i>	<i>the status quo</i>

The *Cambridge* dictionary also contains a non-homogeneous set of compound words, such as these examples in Table 18.

Table 18 Some *Cambridge* hyphenated words not meeting the core idiom criteria

<i>argy-bargy</i>	<i>fat-cat</i>	<i>likety-split</i>
<i>a balls-up</i>	<i>a fuddy-duddy</i>	<i>the nitty-gritty</i>
<i>a blue-eyed boy</i>	<i>elbow grease</i>	<i>riff-raff</i>
<i>deep-six</i>	<i>stalking horse</i>	<i>stool pigeon</i>

Many of these (*argy-bargy*, *fuddy-duddy*, *nitty-gritty*, *riff-raff*) are what can be termed ‘phonologically motivated compounds’ (Bauer & Huddleston, 2002), or ‘opaque compounds’ (McCarthy, 1992, 1998). This particular group does not meet the criteria because they are comprised of words which do not occur elsewhere.

With regard to compounds, this study is about MWUs. Compounds are defined as single lexemes. If they are single lexemes, they are not in themselves MWUs, and therefore do not meet the criteria. However, there is a problem with many N + N constructions – are they compounds (lexemes) or syntactic structures? Only if the latter will they be relevant to this study of core idioms. While Bauer & Huddleston (2002) distinguish on some orthographic criterion, Bauer (1998) argues that we cannot distinguish between a (N+N) compound and (N+N) syntactic group, and in common with many other morphologists, calls them all compounds. As *Cambridge* includes all types of MWU, ranging from non-homogenous hyphenated compounds (*a balls-up*) to two word N+N (*elbow grease*) and adj+N (*fat-cat*) constructions and does not distinguish between them, compound words do not meet the criteria to qualify as core idioms.

Another category found in *Cambridge* is MWUs comprising names, either personal names or names of fictional characters in literature or on the screen (movie or television) which are figuratives rather than core idioms, as shown in Table 19:

Table 19 Some *Cambridge* personal name and fictional character MWUs not meeting the core idiom criteria

Joe Bloggs, Joe Blow	Joe Public, John Q Public	John Doe, Jane Doe	a Jack the lad
John Bull	Jekyll and Hyde	Mickey-Mouse	Prince Charming

Yet another *Cambridge* category not meeting the criteria are those which are literal, or based on a literal situation, such as shown in Table 20:

Table 20 Some *Cambridge* literal MWUs not meeting the core idiom criteria

<i>take / plead the Fifth (Amendment)</i> (based on ‘pleading the Fifth Amendment’ in the USA, to allow yourself to remain silent so as not to incriminate yourself)	<i>the birds and the bees</i> (based on the difficulty of teaching young people about sex, so using two species far enough removed from humans to be able to teach about procreation without embarrassment)
---	---

Some ‘literal’ MWUs merit further discussion. For example, MWUs such as *a catch* (-) 22, also based on a literal situation (the 1961 novel *Catch 22* by Joseph Heller) use the idea of ‘catch’ in ‘What’s the catch?’, with ‘catch’ having the dictionary meaning of ‘a concealed, unexpected, or unforeseen drawback or handicap’ (*Collins English Dictionary*, Sinclair, 1991b, p.254). In the novel, there were a number of unforeseen ‘catches’. Therefore, this MWU is based on the literal example given in the novel.

In addition to literal MWUs are euphemisms, often talking about uncomfortable subjects such as sex, or going to the toilet, as Table 21 shows:

Table 21 Some *Cambridge* euphemism MWUs not meeting the core idiom criteria

The earth moved	Go to see a man about a dog
-----------------	-----------------------------

With regard to the category of phrasal verbs, while some idiom dictionaries include varying numbers of phrasal verbs, others such as *Cambridge* include none, possibly because *Cambridge* (McCarthy & Walter, 1997), *Oxford* (Cowie & Mackin, 1993) and *Collins COBUILD* (Sinclair & Moon, 1989) have separate phrasal verb dictionaries. However, as discussed previously (see 2.2.2), a decision was made to exclude phrasal verbs from this study as this large and important group of MWUs deserves separate and thorough research on its own. As well as MWUs from *Cambridge* not meeting the core idiom criteria, a number of other categories were found in other idiom dictionaries.

4.12 ADDITIONAL MWUs NOT MEETING THE ‘CORE IDIOMS’ CRITERIA

Some of the same categories from *Cambridge* which do not meet the core idiom criteria (foreign word MWUs, compound words, personal name and fictional character MWUs) were also found in other idiom dictionaries, or in the BNC.

For example, euphemism and literal MWUs found in other idiom dictionaries do not meet the core idiom criteria because they are compositional or based on literal situations, such as shown in Table 22 below:

Table 22 Some literal and euphemism MWUs not meeting the core idiom criteria

<i>More tea, Vicar?</i> (a humorous expression said after someone has said or done something embarrassing or shocking, <i>Longman</i> , p.340)	<i>The sun is over the yard-arm</i> (it is the time of day when it is permissible to drink alcohol, <i>Oxford</i> , 1999, p.346)
--	--

Additional MWUs were also found to not meet the core idiom criteria because they were simply used to strengthen a question (how, what, when, where, why) or statement, including those found in Table 23:

Table 23 Some strengthener MWUs not meeting the core idiom criteria

in the world / on earth	(in) the deuce / dickens / devil / blazes
(in) the heck / hell / bloody hell / fuck	in the devil's / God's / heaven's name
for shit's sake / for fuck's sake	holy hell / holy Jesus / holy shit / holy fuck

Another category which does not meet the core idiom criteria is that of truncated familiar and institutionalised MWUs such as proverbs (see McCarthy & O'Dell, 2002, p.66). These would be figuratives in their full form. Table 24 gives other MWUs also found in their truncated form.

Table 24 Some truncated idiom dictionary MWUs not meeting the core idiom criteria

<i>a bloody shovel</i> (= truncated version of 'call a spade a spade when it's really a bloody shovel', <i>Oxford</i> , Cowie, Mackin & McCaig, 1983/93)	<i>hair of the dog</i> (= truncated version of 'hair of the dog that bit you', <i>Wordsworth</i> , Kirkpatrick & Schwarz, 1993)
<i>it's brass monkeys</i> (= truncated version of 'brass monkey weather', <i>Longman</i> , Stern, 1998)	<i>head over heels</i> (= truncated version of 'head over heels in love', <i>Oxford</i> , Cowie, Mackin & McCaig, 1983/93)
<i>up the creek</i> (= truncated version of 'up the creek without a paddle', <i>Wordsworth</i> , Kirkpatrick & Schwarz, 1993)	<i>Up yours!</i> (= truncated version of 'up your arse/ass', <i>Cambridge</i> , McCarthy & Walter, 1998)

Expressions of 'emotion' do not meet the criteria, either because they contain words which do not occur elsewhere, are euphemisms, or are figuratives and ONCEs. While *Cambridge* included only a small number (*For crying out loud!*, *Stone the crows!*, *Hells bells!*, *My arse!*, *My foot!*), other idiom dictionaries contained more. A sample of these is included in Table 25 below:

Table 25 Some expressions of emotion not meeting the core idiom criteria

<i>My sainted aunt / giddy aunt</i>	<i>Blow/bugger/damn/dear me</i>	<i>You're putting/having me on</i>
<i>Bless me / my soul</i>	<i>Come off it / I'll give you what for</i>	<i>For the love of God / heaven / Mike / Pete</i>
<i>Cor blimey / Gee whiz</i>	<i>Pig's arse / In a pig's ear</i>	<i>Saints alive / preserve us</i>
<i>I'll be blessed / blowed / bugged/damned/jiggered</i>	<i>For cripes / Christ's / God's / goodness / heaven's sake</i>	<i>God / heaven / Lord forbid</i>
<i>Holy cow / Jesus / mackerel / moly / Moses / smoke/Toldeo</i>	<i>By cracky / gad / George / God / golly / gosh / gum /jove</i>	<i>Good God / gracious / grief / heavens / Lord</i>

Other MWUs do not meet the criteria because they are judged to be literary, cultural, biblical or historical, including the examples in Table 26:

Table 26 Some literary, cultural, biblical or historical MWUs not meeting the core idiom criteria

<i>pile/heap Pelion on Ossa</i> : add an extra difficulty, task, etc. to an already difficult situation or undertaking [With reference to the Greek legend of how the giants in their war against the Olympian gods attempted to stack the Thessalian mountains of Ossa on top of each other in order to scale the heavens and overthrow the gods.] (Speake, 1999, p.262)	<i>the Fourth estate</i> : the Press [The three traditional Estates of the Realm – the Crown, the House of Lords, and the House of Commons – are now viewed as having been joined by the Press, regarded as having power in the land equivalent to that of the others] (Speake, 1999, p.143)
<i>Cast one's bread upon the waters</i> : do good without expecting gratitude or immediate reward [With allusion to Ecclesiastes 11:1 'Cast thy bread upon the waters; for thou shalt find it after many days.'] (Speake, 1999, p.46)	<i>a Potemkin village</i> : a sham or unreal thing [Count Potemkin (1739-91), favourite of Empress Catherine II of Russia, reputedly ordered a number of sham villages to be built for the empress's tour of the Crimea in 1787.] (Speake, 1999, p.374)

Table 27 gives some examples of Cockney 'rhyming slang' MWUs not meeting the core idiom criteria because they represent a single word:

Table 27 Some rhyming slang MWUs not meeting the core idiom criteria

<i>Brahms and Liszt</i> (= pissed) (Longman, 1998)	<i>[not a] dicky / dickie bird</i> (= [not a] word) (Oxford, 2001)	<i>tell porkies</i> (= [pork pies] = lies) (Oxford, 2001)
--	--	---

As explained, the above categories do not meet the *core idioms* criteria and were therefore not included on the list or in the corpus search.

4.13 CONCLUSION

Every attempt was made to not only ensure that the search for core idioms was as wide as possible, but also to apply the core idiom criteria as strictly as possible in order to produce a true list for the corpus search. First a source idiom dictionary, *Cambridge*, was selected, and the core idiom criteria were applied. This resulted in an initial list of 70 core idioms and borderlines. This was followed by a careful examination of other idiom dictionaries, which resulted in another 32 core idioms and borderlines being added. Finally, several written and spoken (or written to be spoken) sources were examined and the core idiom criteria were applied, resulting in only another 2 core

idioms or borderlines being added to the initial list of 70 from *Cambridge* and the 32 from other idiom dictionaries. This meant there were now a total of 104 core idioms and borderlines to use for the corpus search.

The following chapter, Chapter 5, provides details of the corpus search and the results of the search.

CHAPTER 5

CONDUCTING THE CORPUS SEARCH AND RESULTS OF THE CORPUS SEARCH

5.1 INTRODUCTION OF THE CORPUS SEARCH

This chapter gives the procedure for the corpus search, plus some of the results of the corpus search and a frequency indication of the items in the list of the core idioms and borderlines. The British National Corpus (BNC) was selected for the search of core idioms and borderlines for a number of reasons, including:

- its size – 100 million words of written (90%) and spoken (10%) English
- its recency – over 93% of the BNC texts are from the period 1985-94, with spoken data no earlier than 1991
- its availability – the BNC was available for research purposes outside of the U.K. in late 2000
- its relevance – because New Zealand has many British settlers, New Zealand English shares many similarities with British English
- its range and representativeness – the BNC is a finite, balanced, sampled corpus made up of two kinds of written texts (imaginative – 20%, informative – 80%) and two kinds of spoken English, conversational (males and females, aged 15-60+, from different social groups and regions) and task-oriented (lectures, company and trade union talks, business meetings, consultations, sermons, parliamentary and legal proceedings, TV/radio broadcasting, speeches, commentaries, etc.)

(Leech, Rayson, & Wilson, 2001, pp.1-4)

Although an argument could be made that the English which our ESL/EFL students are primarily exposed to is New Zealand English, the Wellington Corpus (written and spoken) contains only a million words of each, so was not

considered large enough to provide enough instances of the MWUs being searched for.

This first part of this chapter discusses the search carried out in the BNC, while the second part gives some results of the search. Frequency lists plus a complete list and number of occurrences of all the core idioms are given in Table 50, and in Appendices 1-7, with the full results of the search including sample sentences of both written English (WE) and spoken English (SE), common collocations, and a historical explanation given on the CD attached. Before looking at how the search was conducted and challenges that arose, we will first discuss the search engine used in the BNC.

5.2 USING THE BNC SEARCH ENGINE, 'SARA'

The search software, SARA, gave various options for the corpus search, including:

- Word Query
- Phrase Query
- Part of Speech Query
- Pattern Query (to find variations in verb form, noun form, etc.)
- SGML Query (to find text type)
- Query Builder (to do a more complex search)
- CQL Query (to find a 'sequel' of words within a word span)

The first three search tools (Word Query, Phrase Query, Part of Speech Query) are self-explanatory. The last four will be explained briefly next.

5.2.1 Conducting the initial search

Initially, in a first search for each core idiom, the Phrase Query was used. However, other search tools proved to be more comprehensive and more effective. The Pattern Query, for example, was used to find differences in a pattern of a phrase, in particular variations in the verb or noun. In other words, if you wanted to find examples of “chew|chews|chewed|chewing the fat”, it was easier to use the Pattern Query search tool than to do separate searches for each verb tense of each phrase. The SGML Query was used for finding the text type(s) that the word/phrase occurs in, or in other words, its range.

The Query Builder was a more complex search tool, allowing a search for pattern variations in a verb or noun. It proved to be the most comprehensive way of searching for the ‘verb-article/pronoun-noun’ core idioms such as *kick the bucket* or *pull sb’s leg*. This allowed pattern variations and combinations of variations such as variations in the verb tense (kick|kicks|kicked|kicking + bucket), and variations in active or passive (pull|pulls|pulled|pulling + leg|legs, leg|legs + pull|pulls|pulled|pulling). While this tool gives great scope for wide variation, it also takes a rather long time to complete the search. A faster search tool was CQL (sequel), which allows for “pull*leg/10”, finding all occurrences of “pull his/her/my, etc. leg” within a span of 10 words. Sometimes it proved faster to do several CQL searches compared to doing a single comprehensive Query Builder search, or sometimes CQL could be used to double-check part of the Query Builder search.

The possibility of using a ‘_’ symbol as a wildcard was useful, but not always comprehensive. Therefore, when searching for a core idiom involving a

personal pronoun, the wildcard option would find occurrences of my/your/his/her/our/their/its, but no occurrences with an additional adjective. However, the CQL search tool found these as well as the others, so the CQL tool proved more efficient.

5.2.2 Conducting a second search

In the corpus search, different search tools were used when searching for different core idioms, either for an initial search or for a second search to try to ascertain whether any occurrences had been missed.

The disadvantage of doing a CQL search for every core idiom was the need to weed out any non-idiom MWUs that occurred when the two words fell within a span of ten words. Therefore, when a fixed part of a core idiom could be identified, such as *clappers* (to find occurrences of *like the clappers*) or *cropper* (to find occurrences of *come/comes/came/coming a (___) cropper*), a Word Query proved to be the most efficient means of searching. Care was needed, however, to ensure that all variations were included, such as hyphenated MWUs. Care was also needed when searching for what might appear to be a fixed part of a core idiom, such as *bee's knees* or *cat's whiskers*, to ensure that a search was made for any possible variations, including variations in punctuation (*bees' knees*, *cats whiskers*).

Other challenges that arose during the search will be discussed now, such as finding all possible variations of core idioms.

5.3 POTENTIAL PROBLEMS CONDUCTION A SEARCH IN THE BNC

Because many ‘idioms’ (as many as 40% according to Moon, 1998a) are not fixed or frozen, care had to be taken that the corpus search found all occurrences of core idioms, including any variations in either written English (WE) or spoken English (SE). Table 28 gives some of the potentially difficult search results:

Table 28 Potentially difficult search results

Core idiom	Results of search
<i>at loggerheads</i>	A Phrase Query found 59 occurrences (58 WE, 1 SE) but a Word Query using ‘loggerheads’ found 63 occurrences, two of which were literal and referring to a place called Loggerheads. The other two, however, used <i>loggerheads</i> without <i>at</i> , as in ‘any argument that has reached <i>loggerheads</i> ’ and ‘this peculiar state of <i>loggerheads</i> ’. While ‘loggerheads’ in these two is not a MWU, it was still important to include them to show learners a truncated variation of the idiomatic MWU.
<i>be all in</i>	A phrase such as ‘all in’ is more commonly found as part of a larger phrase or clause (all in all, all in the room, all in one place, etc.), so it was necessary to do separate searches combined with the verb ‘to be’ (<i>am/is/are/was/were/been/being</i> + <i>all in</i>) plus the verbs ‘look’ (<i>look/looks/looked</i> + <i>all in</i>) and ‘feel’ (<i>feel/feels/felt/feeling</i> + <i>all in</i>). The phrase was also found with adverbs such as <i>about all in</i> or <i>absolutely all in</i> .
<i>red herring(s)</i>	The core idioms are listed in all the dictionaries, but always in the singular form. A search for ‘red herring’ found 56 occurrences, while a search for the hyphenated ‘red-herring’ found another 2. However, when the plural ‘red herrings’ was searched for, an additional 26 occurrences were found. Therefore, all core idioms that could occur in the plural had to be searched for in this form as well.

These examples give some indication of the care that had to be taken when conducting the corpus search so that every attempt was made to ensure that occurrences were not missed. Another complication in the corpus search involved occasions where even more of a judgement call was required.

5.4 MAKING A JUDGEMENT CALL

On a number of occasions it was necessary to make a judgement call about the inclusion or exclusion of certain MWUs. These will be discussed next, starting with occurrences of identical sentences.

5.4.1 Identical sentences

Occasionally, two identical sentences were found containing the core idiom being searched for. One explanation for this is that one sentence may have been part of a headline, or a column heading in a newspaper, and the other sentence may have been contained within the article. In such cases, a decision was made to count the examples separately, as two occurrences rather than as one, as learners would encounter the core idiom twice, possibly in different places. This situation did not occur often enough to cause any serious concern.

5.4.2 Core idioms in isolation

Some core idioms occurred by themselves, prior to a sentence. Once again, it appeared that a media device is to use an idiomatic MWU as a headline. The core idiom may have been a ‘play on words’ related to the article rather than to the meaning, such as in *She’s Beside Herself* used as a song title, and an ‘expression of emotion’ such as *Great Scott* referring to articles about someone with a surname Scott, a first name Scott, the name of a field, and so on. These occurrences were included, but identified whenever it was possible to do so.

5.4.3 Literal vs idiomatic

On occasion, corpus examples had to be examined closely to see whether the example was being used idiomatically or literally. For example, when a core

idiom such as *swing the lead* was used, it was necessary to ensure that the occurrence was not being used literally, for example, but was used in the idiomatic sense of ‘to pretend to be ill so that you do not have to work’ (Cambridge, McCarthy & Walter, 1998, p.222).

5.4.4 Different spellings

Some core idioms had different spellings, even a difference in apostrophe (*bee’s knees*, *bees’ knees*; *the cat’s whiskers*, *the cats whiskers*). As they all qualified as core idioms or borderlines, all occurrences were included in the total count.

Many core idioms are spelled either with, or without, a hyphen (see 4.7.2) so care had to be taken that all spellings were included in the search, as well as variations in form (*take the piss* – *a piss-take*). Therefore, the search included all forms to show learners any possible variations.

5.5 CONCLUSION OF THE CORPUS SEARCH

An attempt was made to conduct a thorough search of the BNC in order to ensure that all possible occurrences of the various core idioms and borderlines were found. Frequently, more than one search tool was used to check that all variations had been found, including the following:

- differences in use
- differences in meaning
- differences in spelling
- differences in form

Other situations requiring a ‘judgement call’ included:

- the core idiom used in isolation
- the identical sentence occurring twice
- literal vs idiomatic

Although these various ‘glitches’ were encountered during the corpus search, a solution was generally found to any problem which occurred. While it is still possible that some occurrences of core idioms were missed, it is hoped that the care taken ensured as complete a search as possible. Next, we will look at some of the results of the corpus search.

5.6 THE CORPUS SEARCH RESULTS

Complete results of the corpus search are provided on the attached CD. It contains almost 100 pages, with approximately 40,000 words, of close-typed text giving extensive information on the 104 core idioms and borderlines.

A search of the BNC proves that these 104 core idioms and borderlines are infrequent in occurrence. They are listed from most to least frequent in the table below, and in separate appendices (see pages 157-160). The ones which were also found with a literal occurrence in the BNC are marked with an asterisk (*), while the ones which are judged to be frozen as they occurred with no variation in form in the BNC or in idiom dictionaries are written in **bold**. As shown in Table 29 below, none of these core idioms occurs often enough to merit inclusion in the 5000 more frequent words of English:

Table 29 Most to least frequent core idioms and number of BNC occurrences

Core idioms and borderlines from highest to lowest number of occurrences			
<i>by and large</i> –	487	<i>by and by</i> –	17
<i>so (-) and (-) so</i> –	327	<i>push the boat out*</i> –	16
<i>such (-) and (-) such</i> –	196	<i>gird (up) your loins</i> –	16
<i>out (-) of (-) hand</i> –	141	<i>hell for leather</i> –	15
<i>take the piss</i> –	137	<i>cock (-) a (-) hoop</i> –	15
<i>and what have you</i> –	136	<i>Bob's your uncle</i> –	14
<i>serves sb right</i> –	101	<i>chew the fat*</i> –	14
<i>take sb to task</i> –	92	<i>cold turkey</i> –	14
<i>red herring(s)</i> –	87	<i>kick the bucket</i> –	13
<i>(be) beside yourself</i> –	72	<i>(be) the cat's whiskers</i> –	13
<i>out (-) and (-) out</i> –	72	<i>give sb the bird*</i> –	13
<i>take the mickey</i> –	71	<i>(be) all in*</i> –	13
<i>at loggerheads</i> –	63	<i>French letter(s)</i> –	12
<i>pull sb's leg*</i> –	60	<i>not cut the mustard</i> –	12
<i>touch (-) and (-) go</i> –	53	<i>odds and sods</i> –	12
<i>the Big Apple</i> –	52	<i>like the clappers</i> –	12
<i>cut no ice with sb</i> –	50	<i>Beat it!*</i> –	12
<i>come a cropper (49) / come a purler (1)</i> –	50	<i>Butter wouldn't melt in sb's mouth</i>	11
<i>a piece of cake* (43) / a piece of piss (7)</i> –	50	<i>sweet Fanny Adams</i> –	10
<i>put your foot in it (37) / put your foot in your mouth (11)</i> –	48	<i>to boot*</i> –	10
<i>an axe to grind</i> –	47	<i>knock into a cocked hat</i> –	9
<i>make no bones about it</i> –	44	<i>cook sb's goose</i> –	8
<i>white elephant(s)</i> –	43	<i>do your nut</i> –	8
<i>(all) of a piece</i> –	38	<i>Old Glory</i> –	8
<i>Uncle Sam*</i> –	35	<i>cut up rough</i> –	8
<i>and what not</i> –	33	<i>pop your clogs</i> –	8
<i>go by the board(s)</i> –	32	<i>bang to rights</i> –	8
<i>So long!</i> –	24	<i>at sixes and sevens</i> –	8
<i>Eat your heart out</i> –	23	<i>(be) the bee's knees</i> –	7
<i>not hold a candle to</i> –	23	<i>like nobody's business</i> –	7
<i>(be) off the wall</i> –	22	<i>cut a rug</i> –	6
<i>pull the other one</i> –	19	<i>(be) off your face</i> –	5
<i>trip the light fantastic</i> –	18	<i>get your rocks off</i> –	5
<i>cut a dash</i> –	18	<i>have a bone to pick with sb</i> –	4
<i>to a tee / T</i> –	18	<i>(be) nip and tuck</i> –	4
		<i>shoot the breeze</i> –	4
		<i>heap coals of fire on sb's head</i> –	3
		<i>swing the lead</i> –	3
		<i>(have) bats in the belfry</i> –	3
		<i>sell sb a bill of goods</i> –	2
		<i>splice the mainbrace</i> –	2
		<i>(be) a fine kettle of fish</i> –	2
		<i>set sb/sth by the ears</i> –	2
		<i>take the rise out of sb</i> –	2
		<i>lay an egg</i> –	1
		<i>how's your father</i> –	1
		<i>make sth out of whole cloth</i> –	1
		<i>all gas and gaiters</i> –	1
		<i>make your number</i> –	1
		<i>half seas over</i> –	1
		<i>hop the twig</i> –	1
		<i>all ends up</i> –	1
		<i>Box and Cox</i> –	0
		<i>buy the farm*</i> –	0
		<i>come the raw prawn</i> –	0
		<i>go bung</i> –	0
		<i>the ghost walks</i> –	0
		<i>part brass rags with</i> –	0
		<i>bust sb's chops</i> –	0
		<i>the bitch goddess</i> –	0
		<i>row crosshanded</i> –	0
		<i>clean sb's clock</i> –	0
		<i>tie one on</i> –	0
		<i>eat crow</i> –	0
		<i>lay out in lavender</i> –	0
		<i>over the transom</i> –	0
		<i>to a fare-thee-well</i> –	0
		<i>peg it</i> –	0
		<i>wear the green willow</i> –	0
		-	

Despite the small number of infrequent core idioms found, the aim of the data on the attached CD is for the information to be of use to both language teachers and learners, in the same way that a dictionary is useful. Therefore, information is provided on each core idiom or borderline figurative/ONCE, including frequency, meaning, variations, examples, collocations, source, historical or etymological information, constraints on use and recognition of literal use. Instructions are given at the beginning of the CD to make it as

‘user-friendly’ as possible. Each of the core idiom descriptions is given in the format above, as demonstrated in this chapter (see 5.8). Before giving any results, a decision had to be made on how those results would be listed, and this is explained next.

5.7 DECIDING HOW TO LIST THE RESULTS OF THE CORPUS SEARCH

A comprehensive list of core idioms and borderlines had been compiled. Before beginning the search in earnest, however, it was important to decide how to list the results of the search. It seemed logical to proceed through the alphabetical list of core idioms in a methodical and orderly way, so as to ensure that none was missed. Therefore, the search began with the list of ‘A’ core idioms and proceeded through the alphabet. When trying to decide how best to list the results, idiom dictionaries were consulted but some discrepancies became clear.

5.7.1 How idiom dictionaries list their ‘idioms’

A decision was made to list the results of the search, regardless of the individual frequency of core idioms, in an alphabetical order – as in all dictionaries – to make for easier searching, with frequency lists included in Table 29, and in the appendices. When deciding how to list each core idiom – under the first word of the MWU, or under the noun, adjective, verb, or adverb of each MWU – other idiom dictionaries were consulted. The result was somewhat surprising. While the *Oxford Dictionary of English Idioms* (Cowie, Mackin & McCaig, 1983, 1993) chose to list the MWUs alphabetically under

the first word of each phrase, other dictionaries seemed to list the idiomatic MWU under the ‘headword’ but this headword varied. Some dictionaries, such as the *Collins COBUILD Dictionary of Idioms* (Sinclair & Moon, 1995, p.viii) explain that generally the headword is a noun, so the ‘idiom’ is listed under the first noun of the MWU (*spill the **beans**, rock the **boat***), or if no nouns, under an adjective (*go **easy** on someone, in **black** and **white***). However, they also list their exceptions:

- if the noun varies (*beat your **chest***), the idiom is listed under the verb
- if two similar idioms would end up in different parts of the dictionary (*a fair crack of the whip, crack the whip*), the idiom is listed under the common noun (**whip**)
- if the idiom contains two nouns but the first noun is very general such as ‘end’ or ‘top’, then the idiom is listed under the second noun (*the thin end of the **wedge***)
- similes such as *white as a sheet* and *old as the hills* are always dealt with under their adjectives (**white** and **old**)

Cambridge (McCarthy & Walter, 1998, p.viii) takes another approach, explaining that ‘idioms’ are listed under a ‘keyword’ (with no explanation for how the keyword is chosen), and suggesting dictionary users check the index at the back of the book to find the keyword, which is printed in bold.

The result of differences such as these is that a learner may encounter the same figurative, ONCE, or core idiom in different dictionaries, with the same core idiom listed in *Cambridge* differently than in other idiom dictionaries, as Table 30 shows:

Table 30 How idiom dictionaries list idiomatic MWUs

Core idiom	Listed in <i>Cambridge</i>	Listed in other idiom dictionaries
<i>cook sb’s goose</i>	under G (p.160)	under C (p.76), <i>Oxford</i> , 1999
<i>red herring</i>	under H (p.207)	under R (p.322), <i>Collins COBUILD</i> , 1995
<i>cold turkey</i>	under C (p.75)	under T (p.364), <i>Longman</i> , 1998
<i>by and large</i>	under L (p.220)	under B (p.47), <i>Wordsworth</i> , 1993

Instructions such as those given in the introduction of each idiom dictionary may seem logical to a first-language (L1) speaker but may not seem as clear to a second-language (L2) speaker struggling to come to terms with idiomatic MWUs of any kind. Therefore, a decision was made in this study to simplify things as much as possible. Because learners generally have a reasonable grasp of grammar – especially nouns and verbs – it was first decided to explain to learners that core idioms were listed under the first noun of each MWU, but if no noun – under the verb, next under an adjective and finally under the adverb, as demonstrated here in Table 31:

Table 31 How core idioms are listed on the attached CD

Core idiom	How listed on the CD
<i>make no bones about it</i>	under B, for the first noun ‘bones’
<i>not hold a candle to somebody</i>	under C, for the first noun ‘candle’
<i>a piece of piss</i>	under P, for the first noun ‘piece’
<i>Beat it!</i>	under B, for the verb ‘beat’

The majority of core idioms were relatively easy to determine (*(be) all in* – under ‘all’, *chew the fat* – under ‘fat’), but others were more difficult. The situation was less straightforward, for example, with a number of core idioms involving either two nouns, or those with capital letters. Therefore, the decision was made to give words written with a capital letter priority, and if no capital letter, then to give priority to the first noun of the MWU:

- *sweet Fanny Adams, a French letter* – under ‘Fanny’ and ‘French’
- *Big Apple, Box and Cox* – under ‘Big’, and ‘Box’
- *Old Glory* – under ‘Old’

The same rule was applied to individual letters, written in either upper or lower case:

- *to a T* – under ‘T’

A number of other core idioms involved two nouns with no capital/upper case letter, such as possessives. In these cases, again it seemed logical to put the core idioms under the first noun:

- *the bee's knees* – under 'bee'
- *the cat's whiskers* – under 'cat'

However, where an adjective and noun occurred together, core idioms were listed under the noun rather than the adjective:

- *red herring* – under 'herring'

To reiterate, priority was given to the capital letter word in any core idiom, and to the first noun of possessive core idioms. Priority was given in other core idioms first to the noun, if no noun to the verb, if no verb to the adjective, if no adjective to the adverb. All of this is explained on the attached CD, giving the complete results of the corpus search. Clearly no system is without fault, but it is hoped that this one is relatively easy for learners to follow.

5.8 FORMAT FOR RESULTS

As mentioned, the results of the search of core idioms on the attached CD are likely to be of interest to both English language teachers and learners. For this reason, it was designed to give both the maximum useful information related to each core idiom, covering nine different areas. One of the areas both teachers and learners will be most interested in is frequency, even with these relatively infrequent core idioms. As well as saying whether the core idiom is in the Top 20 (Y or N), the header of each core idiom also gives information on the number found in written English, spoken English and literal examples. The remainder of the 'Format for Results' includes:

the **meaning**, both of the core idiom and any literal meaning.

1. any **variations** of the core idiom, as well as the total number of each core idiom found, given in square brackets at the end [].
2. **examples from the corpus** of the core idioms included.

Where possible, five examples each are taken from Written English (WE), from Spoken English (SE) and literal examples. Conversation from examples that are from written or scripted English is enclosed in single quotation marks, while conversation from spoken English has no quotation marks and is verbatim, containing hesitations, repetitions, slips, and other features typical of natural conversational English.

3. typical **collocations** of the core idioms.
4. the **source** of the examples, stating which part of the corpus (books, newspapers, board meeting, conversation, and so on) examples are from.
5. the **etymology/historical background** giving any known historical/etymological information, taken from eight different books (listed at the beginning of the CD).
6. the **pattern** which should be learnt first (the most common pattern found in the corpus).
7. some **constraints on use**, mostly related to how often or where the core idiom is used.
8. some **clues to recognise literal use** which either explains what literal use is likely to include, or indicates where there is no literal use.

The aim is for this format to offer both language teachers and learners the sort of information that would be useful to them. As mentioned, when deciding what vocabulary to teach/learn first, frequency is one of the most important factors. For this reason, information on whether the core idiom is in the Top 20 is included, as well as the total number of occurrences of each one, as these examples show.

5.8.1 Examples of the ‘Format for Results’

Three examples of the complete format for the results follow, using core idioms taken from the beginning (A), middle (G) and end (S) of the alphabet:

Core idioms (CI)	Top 20	Frequency in BNC		
		No.& % in WE	No.& % in SE	No.& % literal (L)
ALL	N	11 (85%)	2 (15%)	0 (0%)
<i>be all in</i>				

1. Meaning: CI (to be very tired and unable to do any more)
L (no literal meaning)

2. Variations: *am* (1), *is* (1), *are* (5), *was* (1), *were* (0), *been* (0), *being* (0), *look* (1), *looks* (0), *looked* (1), *looking* (0), *about* (2), *absolutely* (1) *all in* [13]

3. Examples from corpus:

WE: 'Thanks, Peggy, but I'm all in; I think I'll go to bed.'

Immediately he slipped his arm about her. 'You're all in. But it's not much further now...'

'Sit down, you look all in.'

'It's heavy,' she said. 'I'll carry it down for you,' said Philip. She looked all in.

'How are you, my friend?' He smiled wanly. 'About all in.'

'You look exhausted – absolutely all in.'

SE: They're a little bit faded today, we're all in today, which makes it a little bit easier.
Right. Go and sit over there. You're all in then Dad! Stay there love.

Literal: None

4. Typical collocations: to be/look + all in

5. Source/origin: books [WE], conversations [SE]

6. Etymology/Historical background: [Slang; second half of 1800s] (Ammer 1997, p.10)
A nineteenth-century Stock Exchange expression describing a market that was depressed because cautious traders were holding *all stock in* until they could see the way ahead. (Donald 1994, p.6)

7. First pattern to learn: *to be all in*

8. Constraints on use: Not used often.

9. Clues to recognise literal use: No literal use with 'all in' by itself.

Core idioms (CI)	Top 20	Frequency in BNC		
		No. & % in WE	No. & % in SE	No. & % literal (L)
GOOSE	N	8 (100%)	0 (0%)	0 (0%)
<i>cook somebody's goose</i>				

1. Meaning: CI (if you cook someone's goose, you do something that spoils their plans and prevents them from succeeding)
L (cook a goose that belongs to someone else)

2. Variations: *cook* (5), *cooks* (0), *cooked* (0), *cooking* (0) ___'s goose, *cook the golden goose* (1); *cook-your-geese*, *goose-cooking/cookers* (0), *somebody's goose is cooked* (0), *our goose is f---ing cooked* (1), *cook/cooks/cooked/cooking-geese* (0) [8]

3. Examples from corpus:

WE: At the time of the lobby revolt, 'No one could have been in any doubt that if they went over to [the rebel papers'] side, they would cook their goose with Number Ten.'

This will cook Arthur's goose.

i.e. cook ___'s goose, constitutes a minimal semantic constituent...

It has long been recognised that expressions such as to pull someone's leg, to have a bee in one's bonnet, to kick the bucket, to cook someone's goose, to be off one's rocker, round the bend, up the creek, etc. are semantically peculiar.

...unless the commercial manipulators in tennis are careful, they could cook the golden goose. 'Aren't you worried that you'll kill the goose that laid the golden egg?' 'We are the goose,' says Ian, 'and our goose is f---ing cooked.' 'Who served Sir Thomas's goose?' He'd nearly said cooked. There was a silence. James and Algernon glanced at each other. Nothing was said. Auguste came to the rescue. 'Inspector, this is the problem. Anything tampered with in the kitchen could not be guaranteed to reach Sir Thomas unless the poisoner served it himself.'

SE: None

Literal: None

4. Typical collocations: would cook ____'s goose

5. Source/origin: book, magazines [WE]

6. Etymology/Historical background: There are numerous colorful theories about the origin of this term. According to one, the inhabitants of a besieged town in the sixteenth century hung out a goose to show their attackers they were not starving and so enraged them that they set fire to the town and thus cooked the goose. According to another, it comes from the fable about the goose that laid the golden eggs, which, when the farmer killed it to obtain the gold inside, left him with nothing but a goose to cook. The earliest written records of the term date from the mid-nineteenth century, one being in a street ballad opposing the Pope's appointment of a particular cardinal ("If they'll come here we'll cook their goose, the Pope and Cardinal Wiseman," 1851). (Ammer 1992, p.69)

The origin of this phrase has been lost, but there are numerous fanciful theories; one concerns a besieged town that displayed a goose to show it had enough food, causing the attackers to set it on fire. The first recorded use of this colloquial phrase was in 1851. (Ammer 1997, p.135)

The origin of this expression, not noted until the early-mid nineteenth century, is unknown and is perhaps nothing more complicated than an analogy with countless other sayings like 'to settle one's hash'. It could possibly have connections with 'making things hot for someone' perhaps based on 'goose', which at one time signified a goose-neck handled smoothing iron that was heated on the fire. There are also tales of ancient cities that hoisted a goose, the symbol of stupidity, to taunt a besieging enemy, and were subsequently burnt to the ground in reprisal, but the expression can be dated back only to the middle of the nineteenth century. (Donald 1994, pp.78-9)

7. First pattern to learn: *cook ____'s goose*

8. Constraints on use: Used very seldom, but perhaps used more often than occurrences in the BNC indicate.

9. Clues to recognise literal use: Literal use is unlikely, but if someone's goose has been cooked, it should be clear from context.

Core idioms (CI)	Top 20	Frequency in BNC		
		No. & % in WE	No. & % in SE	No. & % literal (L)
SUCH	Y	139 (71%)	57 (29%)	0 (0%)
<i>such and such / such-and-such</i>				

1. Meaning: CI (a particular event, object, etc. that has been specified previously but which is not directly named in a later report)

L (no real literal meaning without the particular event, object, etc.)

2. Variations: *such and such* (61), *such and such a* (64), *such and such an* (4), *such-and-such* (67) [196]

3. Examples from corpus:

WE: EISENHOWER: My military leaders say, 'We need such and such.'

‘Just you ask me where such and such is and I can place it...just like that!’
 Wouldn’t it be terrible if such and such happened?
 Suppose I ask such and such, and why is the floor manager waving his arms about?
 Students are often asked to criticise such and such an author’s views, or to appraise critically the outcome of a particular idea or thought.

SE: God just decides that such and such is correct or such and such is wrong.
 If he asked me such and such I’d say now this is how you do it.
 ...maybe we even have cut-off dates so that if there isn’t, you know, a script by such and such a date then we will need to figure out what these dates were...
 ...the choice is not the same but they should be aware of some of the implications if they chose to go down such and such a road.
 ...I’d rather somebody went along sort of said right I’ll have it done such and such a time rather than keep bugging about.

Literal: None

4. Typical collocations: in/on/at/or/by/is/was/knew/wearing/read/sing/video/be/put/diverted to/do/does/capable of doing/meant/observe/need/discovered/eating/dreaming of + such and such, such and such + a time/a thing/a sort/a way/a compartment/a registration/a fella/a place/a day/a date/a library/a meeting/a coach/a race/a paper/a colour/a dress/a weight/a concept/a position/a distance/a fact/a scholar/a reason/a case/a dream/a police force/a tree/a flower/a brand/a night/a pollutant/an evening/a make/a book/a photographer/a transistor/an interview/a point/a crime/a yard/a trait

5. Source/origin: books, newspapers, magazines, Weekly Hansard, unpublished miscellanea [WE], church report, public meetings, group meeting, business meeting, company meeting, committee meeting, managerial meeting, training session, training courses, job interviews, teachers’ conference, radio broadcasts, radio programme, interviews, conversations [SE]

6. Etymology/Historical background: [Mid-1500s] (Ammer 1997, p.612)

7. First pattern to learn: *such and such, such-and-such*

8. Constraints on use: Used often in both spoken and written English, but especially in spoken English, to refer to something unspecified.

9. Clues to recognise literal use: No literal use.

This same format was repeated for each of the 104 core idioms and borderlines.

Next we will look at the relevance of this to both teachers and learners.

5.8.2 Relevance of the ‘criteria’ to teachers and learners

Teachers need to remind learners of the two criteria used to define a core idiom – non-compositionality and non-figurativeness – at each stage of the learning.

To begin with, core idioms are non-compositional, so:

- learners have to be able to recognise the form of the MWU (where it both begins and ends)

- learners have to look at the core idiom as a whole phrase/MWU, not at its separate elements/words
- the meaning of the individual elements/words that make up a core idiom cannot be depended on to work out the overall meaning of the MWU

Furthermore, core idioms are non-figurative, so:

- there is no ‘figure’ to help learners work out the meaning of the core idiom
- there are no clear shortcuts to remembering the meaning of the whole MWU, it must simply be learnt and remembered

Core idioms, like other new words, will have to be learnt as a whole. Next we will examine the usefulness of the ‘Format for Results’ for both language teachers and learners.

5.8.3 Relevance of ‘Format for Results’ to teachers and learners

With regard to the ‘Format for Results’, the same pattern is repeated for each core idiom. The aim is for the information to be useful to language teachers and learners, as well as to researchers. Table 32 describes the relevance of each of the nine parts to teaching and learning if the need or desire exists:

Table 32 Relevance of ‘Format for Results’ to teacher and/or learners

Name given in ‘Format for Results’	Relevance to teachers and/or learners
1. Meaning	Learners must learn the idiomatic meaning of the core idiom as a whole. Learners must be able to recognise the core idiom and thus be able to distinguish it from any literal meaning of the MWU.
2. Variations / Total number of occurrences	Learners need to know about variations as they may make recognition of core idioms more difficult. Therefore, teachers and learners should be aware of all possible variations, and learners must be able to recognise them as variations of the CI. In addition, the total number of occurrences of the CI is included, so users can see at a glance how frequent each one is.
3. Examples from corpus	Learners can see how the CI is used in context, and can see whether there are any examples of variations or of the literal use of the CI.
4. Typical collocations	Teachers and learners can see which verbs, nouns, etc. most often collocate with the CI. Learners can also see how these combinations of the CI and collocates differ from the literal meaning of the CI and thus help recognise the idiomatic uses.
5. Source / origin	Learners can see in what types of texts the CI can be found – its range of the CI. Teachers can also help them understand how the CI may be used differently in different types of texts so they begin to recognise occurrences when they see them.
6. Etymology / Historical background	Learners can discover the etymological/historical explanation of the CI, which may be a tool in helping learners recognise the MWU and remember its meaning. The explanation may even move the core idiom into the figuratives group by providing the ‘figure’.
7. First pattern to learn	Learners can see which pattern of the CI is the most common, and thus the most useful one to deal with.
8. Constraints on use	Teachers and learners need to know how often the CI is used to help them decide how much time to spend on it.
9. Clues to recognise literal use	Learners need to know whether the literal meaning exists, and if so, how to recognise it and how to distinguish it from the CI meaning.

Therefore, information can be used in the following ways:

- frequency and source can be used to select higher frequency core idioms found in areas learners are likely to encounter them
- where literal meanings occur, the difference between idiomatic and literal meanings can be highlighted
- variations can be examined, to see which ones exist for each core idiom and how common they appear to be, as well as to discuss whether any other variations are possible
- examples of the core idiom used in both written and spoken English can be compared and contrasted to find similarities and differences
- the historical/etymological information in ‘Etymology/Historical background’ might be one of the best ways of helping learners both better understand and remember the core idiom, and may even make the MWU figurative for them

Next we will discuss variations found in the BNC.

5.9 OCCURRENCES INVOLVING VARIATIONS

Like Moon (1998a) who devoted a whole chapter in her book to variations, I found a number of them, including variations in hyphenation, in spelling, and in form, and these will be discussed next, beginning with variations in hyphenation.

5.9.1 Variations in hyphenation

English orthography is extremely inconsistent when dealing with compounds (see 4.11), and as explained previously, although a decision was made in the beginning to exclude compound words, it was found that the hyphenation was often a matter of choice (see 4.7.2), and so frequently the same hyphenated ‘idiom’ was found as a MWU.

Some core idioms, for example, were written in *Cambridge* only in their hyphenated form (*touch-and-go*), while other dictionaries listed them in their non-hyphenated form: *touch and go* (*Collins COBUILD Dictionary of Idioms*, Sinclair & Moon, 1995). Some occurred as frequently or even more frequently, in their non-hyphenated form:

- *touch and go* (46), *touch-and-go* (7)

Others occurred less often in their non-hyphenated form but still added significantly to the total number found:

- *out-and-out* (49), *out and out* (23)

A sample of more of these variations of core idioms is provided in Table 33:

Table 33 A sample of occurrences of hyphenated and non-hyphenated variations

<i>by and by</i>	(15)	<i>by and large</i>	(485)	<i>out and out</i>	(23)
<i>by-and-by</i>	(2)	<i>by-and-large</i>	(2)	<i>out-and-out</i>	(49)
<i>so and so(s)</i>	(209)	<i>such and such</i>	(129)	<i>red herring(s)</i>	(82)
<i>so-and-so(s)</i>	(118)	<i>such-and-such</i>	(67)	<i>red-herring(s)</i>	(5)
<i>cock a hoop</i>	(2)	<i>a/the French letter(s)</i>	(12)	<i>(be all) of a piece</i>	(37)
<i>cock-a-hoop</i>	(13)	<i>a/the French-letter(s)</i>	(1)	<i>(be) all-of-a-piece</i>	(1)

This discovery led to a search for the hyphenated version of a number of core idioms, even when the hyphenated spelling was unexpected, with Table 34 listing examples of some slightly unusual results:

Table 34 Unexpected variations involving hyphenation found in the BNC

<i>(rejecting) out-of-hand</i>	(1)	<i>all-of-a-piece</i>	(1)	<i>odds-and-sods</i>	(1)
<i>the light-fantastic</i>	(1)	<i>at logger-heads</i>	(1)	-	

Compound modifiers are often hyphenated (Trask, 1997), but not when used as nouns rather than as modifiers. However, examples of hyphenated nouns, used as nouns rather than as modifiers, were also found:

- Although it is only a brief summary, it contains all the relevant information necessary to follow the ten steps – plus a measure of *red-herrings*.

Similarly, when all 487 occurrences of the MWU are used as modifiers but only a couple are hyphenated, it strengthens the argument that hyphenation is very much a matter of personal taste:

- 485 occurrences of *by and large*, 2 occurrences of *by-and-large*

As a general rule, most compound adjectives in British English are hyphenated, while those in American English are not (Partridge, 1978), yet the occurrences found in the BNC again support the generalisation that personal preference and reducing ambiguity both play a part. The ‘moral of this lesson’ – as far as the classroom goes – is to make sure that learners are aware of all possible variations of hyphenation related to the same idiomatic MWU. Other variations included variations in spelling.

5.9.2 Variations in spelling

Some differences in spelling in any corpus can be attributed to the mistranscription of the spoken language. Other differences in spelling found in the BNC were as a result of inflectional differences:

- *bang to rights* (7), *banged to rights* (1)

Still other variations in spelling, as indicated by the ones shown in Table 35 below, range from the difference of a single word, to a difference in punctuation (apostrophe and hyphenation), to the difference of one word of the idiomatic MWU being different, or to the difference of the full words or a truncated form.

Table 35 Range of differences in spelling of core idioms

Apostrophe or no apostrophe	Difference of one word	Full version or abbreviated form	Hyphenated or unhyphenated
<i>bee's knees / bees' knees</i>	<i>cut a / the rug</i>	<i>sweet Fanny Adams</i>	<i>what have you /</i>
<i>the cat's pyjamas /</i>	<i>or what have you / and</i>	<i>/ sweet FA</i>	<i>what-have-you</i>
<i>the cats pyjamas</i>	<i>what have you</i>		
<i>the cat's whiskers /</i>	<i>and what not / or what not</i>	<i>to a tee /</i>	<i>what not /</i>
<i>the cats whiskers</i>	<i>/ the what not / a what-not</i>	<i>to a T</i>	<i>what-not</i>
	<i>/ in a colourful what-not</i>		

As well as variations in spelling, there are also variations in singulars and plurals, and these will be examined next.

5.9.3 Variations in singulars/plurals

Learners know that most nouns, with the exception of uncountable ones, can occur in the plural form. It is less easy to generalise about the rule for nouns in core idioms, however. Some core idioms, especially the 'adjective-noun' collocations, follow the 'rule', with countable nouns taking the plural form:

- *red herring – red herrings*

Others with the 'adjective-noun' form do not allow the plural form:

- *a fine/pretty kettle of fish*, not 'fine/pretty kettles of fish'

- *cold turkey*, not ‘cold turkeys’

This might be because their meanings also do not take the plural form:

- *the Big Apple* – New York – no plural form

Or perhaps it is because the noun they represent is already in the plural form:

- *cold turkey* – the unpleasant mental and physical effects that someone suffers when they suddenly stop taking drugs – already in plural form

Many other core idioms are also not found in their plural form:

- *kick the bucket, lay an egg, trip the light fantastic, shoot the breeze, splice the mainbrace, cut the rug, take the mickey, a piece of cake, (be) off the wall, swing the lead, take sb to task, cook sb’s goose*

Still other core idiom nouns are in the plural but do not take the singular form:

- *at loggerheads, odds and sods, pop your clogs, have bats in the belfry, make no bones about it*

Some nouns in core idioms use the definite article, implying that there might only be only one, or a definite one (*the cat’s whiskers, give sb the bird*), but such may not be the case. Other nouns used in core idioms also do not take the plural form because they are non-countable (*a piece of piss*) or are already in their plural form (*odds and sods*). Table 36 gives further examples of the variations described above:

Table 36 Range of core idioms allowing or not allowing the plural form

Countable nouns allowing plural	<i>red herring</i> – <i>red herrings</i>	<i>white elephant</i> – <i>white elephants</i>	<i>eat your heart out</i> – <i>eat your hearts out</i>	<i>go by the board</i> – <i>go by the boards</i>
Countable nouns allowing plural but not found in plural	<i>not hold a candle to sb</i> (not ‘hold candles to somebody’)	<i>push the boat out</i> (not ‘push the boats out’)	<i>nip and tuck</i> (not ‘nips and tucks’)	<i>chew the rag</i> (not ‘chew the rags’)
Countable nouns whose non-idiomatic meaning does not allow plural	<i>a piece of piss</i> – very easy to do (no plural of ‘very easy’)	<i>do your nut</i> – (be) very angry (no plural of ‘very angry’)	<i>to boot</i> – as well, in addition (no plural of either one)	<i>make sth out of whole cloth</i> – not true (no plural of ‘not true’)
Countable plural nouns allowing singular but not found in singular	<i>the cat’s whiskers</i> (not, ‘the cat’s whisker’)	<i>make no bones about it</i> (not ‘make no bone about it’)	<i>bust sb’s chops</i> (not ‘bust sb’s chop’)	<i>(be) the bee’s knees</i> (not ‘(be) the bee’s knee’)

Learners need to understand that rather than core idioms being ‘fixed’ or ‘frozen’ in terms of singulars/plurals, they allow some variations, often in the same way that their non-idiomatic versions do. Some core idioms can also be truncated.

5.9.4 Truncated core idioms

Not all MWUs, such as proverbs, are used in their full form (see McCarthy & O’Dell, 2002, p.66; Simpson & Mendis, 2003, p.436). It appears that there is a process of elision where familiar idiomatic MWUs are used more often in their shortened form (see Watson, 1998, p.18, *silver lining*; Simpson & Mendis, 2003, p.436, *haven’t the foggiest*). Because they are so familiar, some fixed MWUs become truncated because they are predictable. Core idioms, such as these BNC examples, can also be truncated perhaps for the same reasons:

- *Butter wouldn’t melt in somebody’s mouth – Butter wouldn’t melt*
- *Heap coals of fire on somebody’s head – Coals of fire*
- *put your foot in your mouth – put your foot in it*
- *at loggerheads – Loggerheads is too mild a term for it*

This lack of ‘frozenness’ is even more obvious when examining other variations in form, especially lexical form.

5.9.5 Variations in lexical form

As well as differences in spelling (hyphenation), morphology (singulars and plurals) in inflection (particularly verb tenses), and truncation, idiomatic MWUs often vary in their lexical form, making it necessary to search for a variety of possibilities. The possibility of “creative blending” also complicates the issue so that even when learners become familiar with the canonical form, they may also need to be aware of combinations, accidental or deliberate (‘flip

side of the same coin', Simpson & Mendis, 2003, p.436). Some variations found in the BNC include a difference in the verb used in the core idiom:

- *get / pull your finger out*
- *are / look all in*
- *took / held / called / brought to task*
- *be / look like the bee's knees*
- *give / get the bird*
- *go / slip by the board*
- *make / create something out of whole cloth*
- *go / drive / ride / head hell for leather*
- *have / caught / got beat dead to rights*

Some allow a difference in modification:

- *cut no / cut less / cut very little / not cut any / not cut much / cut more / cut considerable ice with somebody*
- *cut the / any jazz-rock / the etymological / the typographical / your own mustard*

Some allow an addition to, or subtraction from, the core idiom:

- *pull the other one – pull the other one, it's got bells on*
- *look as if butter wouldn't melt in his mouth – butter wouldn't melt*
- *neither fish nor fowl nor good red herring – red herring*

Some allow the substitution of one word or one part of the core idiom:

- *(be) a piece of cake – (be) a piece of piss*
- *take the rise out of sb / take a rise out of sb*
- *cut a / the / their own rug*

Some allow passivisation:

- *pull his leg / having his leg pulled*

Some core idioms allow a variety of different variations, as in the example here and the examples in Table 37 below:

- *make ./ makes / made / making no bones about it / never made any bones about it / without making any bones about it / and no bones about it*

Table 37 Variations in individual core idioms or borderlines

<i>nobody / no other who can hold a candle to ____</i>	<i>put her / his / your / my foot in her / his / your / my mouth</i>	<i>an / any / the / another / no /</i>
<i>he can't / cannot / could / couldn't / doesn't hold a candle to ____</i>	<i>putting his foot in his mouth</i>	<i>no particular / a particular /</i>
<i>unable to / not good enough to hold a candle to ____</i>	<i>up to the kneecap</i>	<i>no political / a political or</i>
<i>no one else's will hold a candle to ____</i>	<i>putting his foot well and truly in his mouth</i>	<i>academic / no government</i>
<i>no-one / neither of the Top Yanks holds a candle to ____</i>	<i>with her foot in his mouth</i>	<i>or party / a personal / a</i>
<i>someone with ____ and hold a candle to ____</i>	<i>foot and mouth</i>	<i>special / a different axe to grind</i>
<i>Hold a candle to that</i>	<i>foot and mouth disease</i>	<i>his own / its own</i>
	<i>foot-in-mouth moments</i>	<i>technological axe to grind /</i>
	<i>foot-and-mouth soldiers</i>	<i>some axe or other / axes to grind</i>
	<i>put / puts / putting ____ foot in it</i>	<i>the axe grinders / axe grinding / to grind axes /</i>
	<i>My foot in it. My mouth in it.</i>	<i>grind the axe / grind my</i>
	<i>put your feet in it</i>	<i>own axe / grinding an axe /</i>
	<i>putting your big feet in it</i>	<i>grinding axes / axe-grinding</i>
<i>at loggerheads</i>	<i>take / takes / took / taking /</i>	<i>go / went / going / driven /</i>
<i>state of loggerheads</i>	<i>taken the mickey</i>	<i>riding / head / says hell for</i>
<i>reaches / been at</i>	<i>take / takes / took / taking the</i>	<i>leather</i>
<i>loggerheads / Loggerheads</i>	<i>mick</i>	<i>hell for leather</i>
<i>would be too mild a term</i>	<i>mickey-taking</i>	<i>hell-for-leather</i>

Note that variations can also be used to express the opposite:

- *cut no ice with – cut considerable ice with*

Clearly the more familiar and institutionalised the core idiom becomes, the more variations it allows. Journalism is especially well-known for manipulating idiomatic MWUs for humour or 'cleverness'. As with the hyphenations and the different spellings of the MWU, all the different forms found – if they qualified as a MWU – were included in the total count of each

core idiom. The more variations that are allowed, however, the greater is the challenge to ESL/EFL learners. Both teachers and learners would benefit by knowing which core idioms are completely frozen.

5.9.6 Frozen core idioms

In considering the ‘frozenness’ of a core idiom, it was decided that those core idioms which allow only a verb inflection or a change in pronoun (my, your, his, their, and so on) are still considered frozen. Furthermore, almost any core idiom can be modified, internally or externally, with a ‘strengtheners’ such as a swear word. Therefore, a core idiom such as *like the clappers* occurred once as *like the bloody clappers*, but was still considered frozen.

Another factor to take into consideration is the core idioms which occur as part of a ‘play on words’, such as *bats in the belfry* (found as a book title about clerical life entitled *Bare Butts in the Belfry*, and a story about publishers entitled *Cats in the Belfry*). As a ‘play on words’ can be done with almost any MWU, this was not considered a factor to determine frozenness. Therefore, *bats in the belfry* is still considered frozen.

If, however, the core idiom occurs differently in the BNC from how it is listed in *Cambridge* or other dictionaries, it is not considered frozen:

- *all in* was listed as *be all in* but also occurs with ‘look’
- *go cold turkey* was listed in two idiom dictionaries, and listed as *cold turkey* in another two, but also occurs with another verb (‘come off’)
- *heap coals of fire on somebody’s head* also occurs in its truncated form, *coals of fire*

None of these is considered frozen. In addition, core idioms which allow an internal modifier that is not simply a ‘strengtheners’ (*cut a dash* – *cut an*

affluent fifties dash) are not considered frozen. With these examples in mind, Table 38 lists the 27 core idioms which were not found to occur with any variations (other than a deliberate ‘play on words’ or the addition of a ‘strengtheners’): Not included are those core idioms with zero occurrences.

Table 38 Frozen core idioms

<i>bats in the belfry</i>	<i>pop your clogs</i>	<i>So long</i>
<i>Beat it!</i>	<i>cut up rough</i>	<i>splice the mainbrace</i>
<i>(be) beside yourself</i>	<i>set sb / sth by the ears</i>	<i>nip and tuck</i>
<i>Bob’s your uncle</i>	<i>all ends up</i>	<i>make your number</i>
<i>to boot</i>	<i>(be) off your face</i>	<i>Old Glory</i>
<i>shoot the breeze</i>	<i>how’s your father</i>	<i>get your rocks off</i>
<i>kick the bucket</i>	<i>all gas and gaiters</i>	<i>(be) half seas over</i>
<i>like nobody’s business</i>	<i>knock into a cocked hat</i>	<i>(be) at sixes and sevens</i>
<i>like the clappers</i>	<i>swing the lead</i>	<i>Uncle Sam</i>

If we exclude the core idioms which allow verb inflections, the number of frozen ones reduces to only 17:

- *bats in the belfry, Beat it, beside yourself, Bob’s your uncle, to boot, like nobody’s business, like the clappers, all ends up, off your face, how’s your father, all gas and gaiters, So long, Old Glory, nip and tuck, half seas over, at sixes and sevens, Uncle Sam*

Learners might like to know which core idioms were completely frozen in the BNC and which allowed variations. They also need to be made aware of the possibility of a ‘strengtheners’ being added, plus the journalistic technique of using a ‘play on words’ and how this applies to core idioms / borderlines. Like Moon (1998a), I found only a minority were frozen.

5.10 NUMBER OF VARIATIONS OF CORE IDIOMS

The large number of core idioms allowing variations presents a further challenge to the ESL/EFL learner, who must come to terms with not only the core idioms as seen in the dictionary, but also the many variations they allow.

The number and percentage of the variations for each letter of the alphabet are in Table 39 below. Verb tense inflections (*cook/cooks/cooked/cooking sb's goose*) and changes in pronouns are not counted, nor are a 'play on words'. However, variations such as plurals, hyphenations and truncations are counted. This time the number of core idioms also includes those with zero occurrences.

Table 39 Number and percentage of core idioms / borderlines which allow variations

Core idioms	Allow variations	Don't allow variations	Zero occurrences
A – 2	2 (100%)	0 (0%)	0 (0%)
B – 20	10 (50%)	8 (40%)	2 (10%)
C – 11	5 (46%)	3 (27%)	3 (27%)
D – 1	1 (100%)	0 (0%)	0 (0%)
E – 4	2 (50%)	2 (50%)	0 (0%)
F – 8	4 (50%)	2 (25%)	2 (25%)
G – 5	1 (20%)	1 (20%)	3 (60%)
H – 5	4 (80%)	1 (20%)	0 (0%)
I – 1	1 (100%)	0 (0%)	0 (0%)
J – 0	0 (0%)	0 (0%)	0 (0%)
K – 1	1 (100%)	0 (0%)	0 (0%)
L – 8	5 (63%)	2 (25%)	1 (12%)
M – 3	2 (67%)	1 (33%)	0 (0%)
N – 3	2 (33%)	2 (67%)	0 (0%)
O – 4	3 (75%)	1 (25%)	0 (0%)
P – 5	3 (60%)	0 (0%)	2 (40%)
Q – 0	0 (0%)	0 (0%)	0 (0%)
R – 6	3 (50%)	1 (17%)	2 (33%)
S – 6	4 (67%)	2 (33%)	0 (0%)
T – 6	4 (60%)	0 (0%)	2 (40%)
U – 1	0 (0%)	1 (100%)	0 (0%)
V – 0	0 (0%)	0 (0%)	0 (0%)
W – 4	3 (75%)	0 (0%)	1 (25%)
X – 0	0 (0%)	0 (0%)	0 (0%)
Y – 0	0 (0%)	0 (0%)	0 (0%)
Z – 0	0 (0%)	0 (0%)	0 (0%)

Overall, of the 104 core idioms, excluding the 17 with zero occurrences, 59 (57%) of them allow variations ranging from minor variations in spelling, hyphenation, singulars/plurals, to more major variations including ellipsis/truncations, internal modification, and substitution – in other words, modification of form of one sort or another. Once again, excluding the 17 (16%) with zero occurrences, the remaining 28 (27%) do not allow variations. Clearly, if frozenness of form was included as a criterion for core idioms, the number would be smaller still (see 5.9.6).

One way for language teachers to help learners better understand MWUs such as core idioms is to provide examples of them and their collocations.

5.11 COLLOCATIONS AND CORE IDIOMS

Each core idiom must be learnt as a ‘whole’ rather than as individual words, so the vocabulary learning theory of ‘chunking’ is relevant. This could be to learners advantage as they can overcome the limitation on the amount of information processed and learnt by ‘chunking’ it (Nation, 2001, p.319). In addition, as explained, learners must face the challenge of knowing not only the form of the core idioms, but also any possible variations of this form. Therefore, rather than learning core idioms on their own (just as learning vocabulary in a list), there are advantages for learners in seeing which words commonly collocate with the various core idioms. Linguists have argued the need for learning collocations, as Nation (2001, p.318) notes:

- Language knowledge is collocational knowledge (N.Ellis, 2001)
- All fluent and appropriate language use requires collocational knowledge (Pawley & Syder, 1983)
- Many word parts are used in a limited set of collocations and knowing these is part of what is involved in knowing the words [or, in this case, the core idioms] (Nation, 2001)

Learning which particular words – especially, nouns, verbs, adjectives, adverbs – commonly collocate with core idioms would give learners a better understanding of predictability with regard to core idiom collocations. Other advantages of learning MWUs like core idioms as ‘chunks’ or units have been outlined (see, for example, Nattinger & de Carrico, 1992; Pawley & Syder, 1983; Lewis, 1993, 2000), as have the advantages of learning these even larger

‘chunks’ (see Pawley & Syder’s 1983 ‘lexicalised sentence stems’), such as core idioms plus their collocations:

- it will help learners to ‘chunk’ larger stretches of language, thus reducing processing time as well as increasing fluency with greater speed in their understanding and their speaking
- it will improve learners’ grammatical accuracy
- it will give learners a better understanding of the core idiom when they see which nouns/verbs/adjectives/adverbs it collocates with
- it will help learners gain confidence in both understanding and using the core idiom accurately

A number of core idioms, and their common collocations have been provided as examples, so that language teachers and learners can become familiar with the words that core idioms commonly collocate with.

5.11.1 Noun core idiom collocations

Many core idioms are noun phrases, found in collocations with verbs, and mainly the verb ‘to be’, such as those shown here in Table 40:

Table 40 Noun core idioms which collocate mainly with the verb ‘to be’

<i>The bee’s knees</i> plus verb collocations	<i>Beside himself/herself/ yourself</i> plus verb collocations	<i>The cat’s whiskers</i> plus verb collocations	<i>A piece of cake</i> plus verb collocations
am / is / are	am / is / are	is	is / was
was / were	was / were	was	had all been
been / being	will be / must be / said to be / been / being	were	should be / would be / will be
looking like	had driven him / sent him back / seems almost	-	-

Some noun core idioms collocate with a large number of verbs, both verbs with a similar meaning and verbs with different meanings:

- *red herring(s)*: is, are, was, were, could be, must have been, proved to be, turned out to be

- *red herring(s)*: strives to ignore, be diverted by, seemed, was becoming, described as, dismissed as, did not even exist as, try to sift out, packed full of, can act as, throw in, disposed quickly of, discounted them as, creating

Still other examples are of the nouns that collocate with particular core idioms like (*go/goes/went/gone*) *by the board(s)*, as shown here in Table 41:

Table 41 Nouns/pronouns which collocate with (*go/goes/went/gone*) *by the board(s)*

Noun collocations with <i>___ by the board(s)</i>				
health	belonging	my promotion	all that	education and transport
golf	that work	moral standards	something that	his original platonic outlook
efforts	standards	their principles	rules of politeness	the other two
one	manners	minority services	these considerations	sex education or moral instruction
it ('s)	loyalty	other niceties	regulation dress	those cherished dreams
truth	thirty years	musical values	the wildlife element	scientific batsmanship

An analysis of the typical collocations may help learners notice that many are related to 'standards', especially in morals/principles/ education or manners/politeness/loyalty while others are related to something specific (golf, work, promotion) which has been abandoned. Learners could then think of things in their lives which have *gone by the board(s)*. Similar comparisons with verb core idiom collocations might also be helpful.

5.11.2 Verb core idiom collocations

Some core idioms are used as verbs and collocate with a variety of nouns/pronouns. Table 42 shows four which begin with a verb, plus their noun/pronoun collocations:

Table 42 Nouns/pronouns which collocate with four core idioms beginning with a verb

Noun / pronoun collocations with <i>chew the fat</i>	Noun / pronoun collocations with <i>push the boat out</i>	Noun / pronoun collocations with <i>make no bones about it</i>	Noun /pronoun collocations with <i>pull(ing) sb's leg</i>
Mr Grade	everyone / they / we	Jim Lewis / Miki	you ('re)
Former leaders	Sir James	he / we / I	he ('s) / she ('s)
Selwyn	Paul Lexington/Mrs Y	Freud	the editor's
we	The Franks	The French	they

Learners may wish to practise by substituting other nouns/pronouns until they understand how these all of them can be used.

Some adverb core idioms, such as *by and by* and *by and large*, collocate with a variety of verbs, as shown in Table 43:

Table 43 Verbs which collocate with *by and by* and *by and large*

Verb collocations with <i>by and by</i>		Verb collocations with <i>by and large</i>		
would drop down to the ground	expecting	do not believe	am/is/are/was/were/will be	is/are/was/would be/should be
stopped the car	should be	fell to	has/have/had	notice/noticed
am resolved to bring / quite cowed	shall hear them	do not tackle	manages/managed	succeed / have succeeded
hope to do better	will night come / will become	has used it / recall / know	are grown / are consistent	have worked / is dealt with
became quite dark	will begin to	do a good job	agree / appears	think/thought

Learners might appreciate knowing that the verb ‘to be’ is the most common collocate of both, but that *by and by* can also collocate with any number of verbs as can its non-idiomatic synonym ‘soon’, while *by and large* can collocate with an even larger number of verbs because of its higher frequency.

Table 44 shows another adverb core idiom which collocates with a large number of verbs that share a pattern:

Table 44 Verbs which collocate with *out of hand*

Verb collocations with <i>out of hand</i>			
reject / rejects / rejected / rejecting	dismiss / dismisses / dismissed / dismissing	turned down	kill / killing
came	murder	deal with	slaying
discard	kicked	condemn / condemned	hanged / hanging
damns	sack	defeated	arrested
declining	shot	became	take

Learners may also need reminding that despite this variety, verbs such as reject/dismiss/condemn are the most common collocations with *out of hand*. A number of core idioms are either used as adjectives or collocate with them.

5.11.3 Adjective core idiom collocations

Core idioms are often either used as adjectives, or modified by adjectives.

Some are noun phrases and used to modify a variety of nouns, or used as noun phrases and modified by a variety of adjectives or noun phrases such as these

BNC examples of *white elephant*:

- *white elephant*: a *white elephant* office block/development/parade/stall
- *white elephant*: a debating chamber/complacen/*this picturesque white elephant*

However, like its non-idiomatic synonym, the adjective core idiom *out and out* collocates with a large number of nouns, as examples in Table 45 indicate:

Table 45 Nouns which collocate with *out and out*

Adjective core idiom, <i>out and out</i> with noun collocations			
lie / liar	Tory	atheist	blackguard
fashion fanatic	dementia	objective	cheapskate
conveyance	criminals	mass annihilation	product
freedom	British people	privatisers	win / winner
favourite	thieving	reformers	sports route
screaming	fight	conservationists	all-round entertainer
performance	cheating	joy	defenders
systems integrator	rogue	evangelist	chauvinist
swine	barbarism	fascist state	touring aircraft
mongrel	thug	jealousy	competition
sinner	bigot	flyers	war

Understanding the sorts of nouns used with adjective core idioms may help learners, as might Table 46 with examples of the sorts of descriptors – adjectives and articles – which can collocate with noun core idioms:

Table 46 Descriptors which collocate with *red herring* and *white elephant*

<i>Red herring(s)</i> and some adjective collocations		<i>White elephant(s)</i> and some adjective collocations	
a / the	a bit of a	a / the / that	a military
this	a perfect	a complacent	a picturesque
one	a neat little	a huge / vast / giant/ enormous	a possible / complete
a complete	an immense and vile	an expensive	their annual
one big	an absolute	not another	the debating chamber
another	a philosophical	an under-used	the big white Unix

Adjectives can also be used internally in the core idiom, as shown here in

Table 47:

Table 47 Internal modification of *cut __ ice with*, *cut __ dash*, and *come __ cropper*

Adjectives used internally with <i>cut __ ice with</i>		Adjectives used internally with <i>cut a __ dash</i>	Adjectives used internally with <i>come __ cropper</i>	
<i>cut no ice with</i>	<i>cut more ice with</i>	<i>cut a greater dash</i>	<i>come nasty croppers</i>	<i>come a nasty cropper</i>
<i>cut little/very little ice with</i>	<i>cut considerable ice with</i>	<i>cut an affluent fifties dash</i>	<i>come a catastrophic cropper</i>	<i>come an almighty cropper</i>
<i>cut less ice with</i>	<i>cuts any ice with</i>	<i>cut a particular dash</i>	<i>come a right bloody cropper</i>	<i>come the most frightful cropper</i>
<i>(not) cut any ice with</i>	<i>(not) cut much ice with</i>	<i>cut less of a dash</i>	<i>come such a cropper</i>	<i>come the most appalling cropper</i>

An examination of these may help learners see how many adjectives in each table describe size or quantity, how many describe type, how many are positive, how many are negative, and so on. Learners may also like to see examples of core idioms used as adverbs or modified by them.

5.11.4 Adverb core idiom collocations

Once again, there are both adverb core idioms and adverbs used to describe core idioms. Table 48 shows two core idioms used as adverbs to modify mainly verbs associated with a similar theme:

Table 48 Verbs which collocate with *like the clappers*, *to a T* and *like nobody's business*

<i>Like the clappers</i> and some verb collocations		<i>To a T</i> and some verb collocations		<i>Like nobody's business</i> and some verb collocations	
rode	ran off	done	have us down	ride to hounds	shift paperbacks
grew	work	fits	fitted their bill	get along	twist an ankle
racing	come	suits	summed up	play around	are in and out
going	coming up	sets off x	is our Eileen	-	-

For example, *like the clappers* and *like nobody's business* collocate with verbs of action or movement such as ‘rode / grew / racing / ran / twist / play’, while *to a T* collocates with verbs of description like ‘suits / fits / done’.

Other core idioms are described by adverbs, sometimes internally, to show ‘how’ or to what degree the condition applies:

- *absolutely all in*
- *serve her damn well right*

Learners may need more examples like this, and also need opportunities to experiment with other adverbs to see the possibilities, including which ones allow internal modification. The hope is that all of these examples will aid language learners in their recognition, understanding and appropriate use of core idioms.

Learners also need to know how to distinguish the idiomatic meaning from the literal meaning of the idiomatic MWUs which have both.

5.12 DISTINGUISHING BETWEEN CORE IDIOMS AND LITERAL MWUs

Learning to distinguish the literal MWU from the core idiom is further complicated by polysemy (Laufer, 1990) where the MWU can have several literal meanings. The potential for causing confusion to learners is clearly demonstrated with BNC examples of one core idiom, *Beat it!* which had only 12 core idiom occurrences but 78 literal occurrences including the following:

- Beat it (strike repeatedly – 30) –
 - ...and whenever we came to a thicket or undergrowth we beat it with our sticks to stir up the prey.
 - ‘Tell me what day this is, Eleanor! You tell me or I’ll beat it out of you!’
 - I waded over to it, trying to drag the four hundred weight on its side, and when all else failed, I beat it with my fist...
- Beat it (overcome, surpass, win – 42)
 - ‘Nothing to beat it for shifting the wind!’

- e. Mark made 103 m.p.h. this time and Robinson just failed to beat it at 102 m.p.h., leaving the large crowd breathless!
- f. ‘Can you beat it, Clive? They’re actually running a bet on who I’m going to marry. Isn’t life rich?’
- Beat it (stir vigorously – 6)
 - g. ‘Take a quart of green gooseberries boil them and put them thro’ a sieve, take the while of 3 eggs, beat them to a Froth, put it to the Gooseberries and beat it both together till it looks white...
 - h. ‘Take as much lean of boiled ham as you please, and half the quantity of fat, cut it as thin as possible, beat it very fine with a mortar...’
 - i. ‘...press out the juice, and draw the milk out of the cow’s udder into it, sweeten it with a little sugar, and beat it well with birchen twigs till...’

Learners must first decide if the MWU has an idiomatic or a literal meaning, and if literal, which literal meaning is correct. They must also learn to look at the MWU and the words used with it. For example, all BNC occurrences of *kick the bucket* were core idioms as all literal occurrences included ‘kick the bucket over’.

5.13 RELEVANCE OF THE INFREQUENCY OF CORE IDIOMS

As shown (see Table 29), core idioms are not frequent. Despite this, learners may appreciate knowing that almost half of the 104 core idioms and borderlines (51, or 49%) would not have literal equivalents, either because the literal equivalent is not grammatically correct, or because it is not logical, as Table 49 shows:

Table 49 Core idioms and borderlines where literal equivalents are not possible

Core idioms and borderlines with no possible literal equivalents		
<i>(be) the bee's knees</i>	<i>the ghost walks</i>	<i>(sth is) a piece of piss</i>
<i>(he/she is) beside him/herself</i>	<i>I'll give you what for</i>	<i>come the raw prawn</i>
<i>(a bit of) how's your father</i>	<i>go bung</i>	<i>part brass rags with</i>
<i>make no bones about it</i>	<i>out (-) of (-) hand</i>	<i>(have sb) bang/ dead to rights</i>
<i>shoot the breeze</i>	<i>knock sb/sth into a cocked hat</i>	<i>take the rise out of sb</i>
<i>Butter wouldn't melt in sb's mouth</i>	<i>Sb, eat your heart out!</i>	<i>(be) half seas over</i>
<i>by and by</i>	<i>(go) hell for leather</i>	<i>(be) at sixes and sevens</i>
<i>by and large</i>	<i>red herring</i>	<i>(he/she) pegs it</i>
<i>(be) the cat's whiskers / pyjamas / miaow</i>	<i>(be) cock (-) a (-) hoop</i>	<i>so (-) and (-) so</i>
<i>bust sb's chops</i>	<i>cut no ice with sb</i>	<i>such (-) and (-) such</i>
<i>like the clappers</i>	<i>trip the light fantastic</i>	<i>to a tee / T</i>
<i>pop your clogs</i>	<i>splice the mainbrace</i>	<i>take sb to task</i>
<i>come a cropper / purler</i>	<i>(not) cut the mustard</i>	<i>(be) touch and go</i>
<i>cut a dash</i>	<i>(he/she/it is) nip and tuck</i>	<i>(go / do sth) cold turkey</i>
<i>(he/she) lays an egg</i>	<i>do your nut / bun / prunes</i>	<i>and what have you</i>
<i>to a fare-thee-well</i>	<i>odds and sods</i>	<i>and what not</i>
<i>put your foot in your mouth</i>	<i>(be) (all) of a piece (with)</i>	<i>wear the green willow</i>

Of the other 53 core idioms and borderlines, an additional 27 (25%) could have possible – but unlikely – literal equivalents (*have a bone to pick with sb*, *heap coals of fire (up) on sb's head*, *a fine kettle of fish*, *have an axe to grind*, *row crosshanded*, *take the piss*). With the remaining 26 core idioms and borderlines, clues would help learners distinguish the literal meaning, such as:

- *Beat it!* – when used as a command (in speaking) or with an exclamation mark (in writing) is likely to be a core idiom. The literal meaning will always include what to beat (eggs, cream, a carpet, a drum, and so on)
- *To boot* – when used with no object that can be booted (a ball) will almost certainly be a core idiom
- *At loggerheads* – when ‘loggerheads’ is used with a small ‘l’, it is a core idiom, but when used with a capital ‘L’, is used literally for the name of a place (1 occurrence)
- *So long* – when used with a comma or fullstop is a core idiom, but when used literally will be used with verbs (be/give/take/wait/get/stay/work), prepositions, conjunctions, or pronouns (for/after/out/and/me/you/him/her/us/them/it).

Learners can see the number of both idiomatic and literal occurrences of these 26 core idioms, given here in Table 50:

Table 50 Number of occurrences of core idioms and their literal equivalents in the BNC

Core idiom	No. of idiom. occur.	No. of literal occur.	Core idiom	No. of idiom. occur.	No. of literal occur.
<i>(be) all in</i>	13	7	<i>a French letter</i>	12	0
<i>the Big Apple</i>	52	0	<i>swing the lead</i>	3	0
<i>Beat it!</i>	12	78	<i>pull sb's leg</i>	60	3
<i>give sb / get the bird</i>	13	7	<i>gird (up) your loins</i>	16	3
<i>push the boat out</i>	16	3	<i>So long</i>	24	0
<i>Bob's your uncle</i>	14	0	<i>take the mickey</i>	71	0
<i>to boot</i>	10	25	<i>out and out</i>	72	6
<i>like nobody's business</i>	7	0	<i>(be) a piece of cake</i>	43	61
<i>make sth out of whole cloth</i>	1	0	<i>cut a / the rug</i>	6	0
<i>(be) off your face</i>	5	0	<i>serve sb right</i>	101	0
<i>sweet Fanny Adams</i>	10	0	<i>cold turkey</i>	14	1
<i>buy the farm</i>	0	2	<i>hop the twig</i>	1	0
<i>chew the fat</i>	14	1	<i>Uncle Sam</i>	35	13

The corpus search reveals that in fact only 13 (about 13%) out of the total of 104 core idioms were found to have literal occurrences, and only 3 (about 3%) had a much higher number of literal occurrences than idiomatic ones. This is significant for learners who should know that the literal occurrences of these MWUs are unusual. Furthermore, the MWUs with the highest number of literal occurrences (*Beat it!*, *a piece of cake*, *to boot*, *Uncle Sam*), should be clearly understood in context. Learners can therefore assume first that the meaning is idiomatic, and only when the idiomatic meaning fails do they need to apply the literal meaning. Next we will look at some of the frequency results of the corpus search of core idioms.

Of the total of 104 core idioms and borderlines, the following was noted:

- 17 of them (16%) had no occurrences in the BNC
- almost half of them (49, or 47%) had fewer than 10 occurrences
- 13 of them (13%) had literal occurrences, including one that had zero idiomatic occurrences in the BNC (*buy the farm*).

Differences exist in occurrences in written and spoken English. A search of the Oxford Hector Pilot Corpus (OHPC) found that unexpectedly ‘pure idioms’ were more common in written English than in spoken (Moon, 1998a, p.309). Furthermore, research on an academic corpus (Simpson & Mendis, 2003, p.425) revealed that two assumptions were unfounded:

- that interactive transcripts would include more idioms than monologic ones
- that more idioms would be found in humanities and social sciences than in the hard sciences

While acknowledging that differences in occurrence in spoken and written English are not simply to do with differences in the sizes or types of the sub-corpora, we can still make some generalisations, written and spoken occurrences in the BNC:

- more than a third of the core idioms and borderlines (40, or 38%) had no occurrences in spoken English
- others were more common in spoken English or approximately equally common in both spoken and written English

Table 51 has examples of this:

Table 51 Examples of BNC core idiom frequency in written and spoken English

Core idioms / borderlines more common in written English	Core idioms / borderlines more common in spoken English		Core idioms / borderlines about equally common in both
<i>(be) beside yourself</i>	<i>(have) an axe to grind</i>	<i>have a bone to pick with sb</i>	<i>(be) at sixes and sevens</i>
<i>come a cropper</i>	<i>by and large</i>	<i>do your nut</i>	<i>(be) the cat's whiskers</i>
<i>out of hand</i>	<i>put our foot in your mouth / in it</i>	<i>like nobody's business</i>	<i>(be) the bee's knees</i>
<i>(be) at loggerheads</i>	<i>take the mickey</i>	<i>like the clappers</i>	<i>chew the fat</i>
<i>(be) a piece of cake</i>	<i>take the piss</i>	<i>go by the board(s)</i>	<i>red herring(s)</i>
<i>take sb to task</i>	<i>serve sb right</i>	<i>Bob's our uncle</i>	<i>pull sb's leg</i>
<i>(be) touch and go</i>	<i>so and so</i>	<i>to a T / tee</i>	<i>out and out</i>
<i>a white elephant</i>	<i>such and such</i>	<i>kick the bucket</i>	<i>make no bones about it</i>
<i>cut no ice with sb</i>	<i>and what have you</i>	<i>and what not</i>	-
<i>trip the light fantastic</i>	<i>pop your clogs</i>	<i>(be) off your face</i>	-
<i>pull the other one (it's got bells on)</i>	<i>a French letter</i>	<i>knock into a cocked hat</i>	-
<i>(be) all of a piece</i>	<i>odds and sods</i>	<i>cold turkey</i>	-

Corpora evidence such as all of the above – despite its limitations – provides examples of useful ‘real life language’ to learners, and allows them to compare it to the language they are learning.

As shown previously (see Table 29), and somewhat surprisingly, core idioms are not frequent in the BNC. With hindsight, it might have been better to have focussed the corpus search on the larger and more frequent ‘figuratives’ category. However as their frequency of core idioms was tested only on the large BNC, it seemed useful to also test the frequency of this list on other – in this case, local, corpora – the WC (Wellington Written Corpus) and WSC (Wellington Spoken Corpus) to see whether the frequency figures remain the same. Although claims have been made that ‘idioms’ are used frequently in N.Z. English (Waretini, 2002), a search of the local Wellington corpora seemed useful to get an indication of how often this category of core idioms are used.

5.13.1 Top 20 core idioms from the BNC compared to the WC / WSC

The top 20 core idioms are listed separately in Appendix 1. Table 52 below compares the numbers of the first 20 of each found in the BNC, as well as in the Wellington Written Corpus (WC) and the Wellington Spoken Corpus (WSC). The BNC is a large corpus comprised of:

- 90 million words of written English
- 10 million words of spoken English

The Wellington corpora are much smaller, containing only:

- 1 million words of written English (WC)
- 1 million words of spoken English (WSC)

Another factor which might explain the different results is that the spoken corpus represents only 10% of the BNC but 50% of the Wellington corpora. Because of the difference in size, the number found in the WC would have to be multiplied by 90, and in the WSC, multiplied by 10, to give equivalent figures to the BNC:

Table 52 The 20 most frequent core idioms in the BNC compared to the WC / WSC

Top 25 core idioms	Number in BNC (WE/SE)	Number in WC (x 90 = BNC number)	Number in WSC (x 10 = BNC number)	Total WC/WSC (WE/SE)
<i>by and large</i>	487	8 (720)	6 (60)	780
<i>so (-) and (-) so</i>	327	2 (180)	13 (130)	310
<i>such (-) and (-) such</i>	196	0 (0)	7 (70)	70
<i>out (-) of (-) hand</i>	141	6 (540)	2 (20)	560
<i>take the piss</i>	137	0 (0)	3 (30)	30
<i>and what have you</i>	136	0 (0)	10 (10)	10
<i>serve sb right</i>	101	1 (90)	2 (20)	110
<i>take someone to task</i>	92	1 (90)	0 (0)	90
<i>red herring(s)</i>	87	1 (90)	0 (0)	90
<i>(be) beside yourself</i>	72	0 (0)	1 (10)	10
<i>out (-) and (-) out</i>	72	0 (0)	0 (0)	0
<i>take the mickey</i>	71	0 (0)	1 (10)	10
<i>(be) at loggerheads</i>	63	1 (90)	0 (0)	90
<i>pull someone's leg</i>	60	2 (180)	2 (20)	200
<i>(be) touch (-) and (-) go</i>	53	0 (0)	4 (40)	40
<i>the Big Apple</i>	52	0 (0)	0 (0)	0
<i>cut no ice with sb</i>	50	0 (0)	0 (0)	0
<i>(be) a piece of cake/piss</i>	50	2 (180)	0 (0)	180
<i>come a cropper(s)</i>	49	1 (90)	0 (0)	90
<i>put your foot in it / in your mouth</i>	48	0 (0)	0 (0)	0

Not much weight can be placed on the comparisons where only 1 or 2 appear in the Wellington Corpora. Comparing the very large BNC with the much smaller WC/WSC might be a bit like comparing the proverbial ‘apples to oranges’, but it gives some indication and allows a few generalisations, including the following:

- *by and large* is used frequently in both corpora
- core idioms including *so and so*, *serve sb right*, *take sb to task*, *red herring*, and *touch and go* would have a similar number of occurrences in both if the corpora size were the same
- core idioms including *such and such*, *take the piss*, *out and out*, *put your foot in your mouth* and *have an axe to grind* were comparatively more frequent on the larger BNC corpus
- core idioms including *out of hand*, *pull sb’s leg*, *(be) a piece of cake/piss*, *come a cropper* and *make no bones about it* were comparatively more frequent on the smaller WC/WSC corpus

While such a comparison may not give accurate numbers, it may prove useful for learners to see which ones are more frequent in the culture in which they are living and studying. In addition, it might help them when they attempt to prioritise the vocabulary they need to learn. With this in mind, it would be useful to have an understanding of the frequency of the other two categories of idiomatic MWUs – figuratives and ONCEs, which we will look at next.

5.14 FREQUENCY OF FIGURATIVES AND ONCEs

Language teachers usually prefer to teach high frequency vocabulary first. However, prior to the availability of corpora evidence, both teachers and textbook writers had to rely on intuitions about the frequency and/or familiarity of certain ‘idioms’ (Popiel & McRae, 1988). Moreover, much depended, and

still depends, on the definition of an ‘idiom’ being used. As shown, idiom dictionaries vary in the number they contain, and this number is not always accurate (see 4.3). Previous estimates of ‘idiom’ frequency, based on different corpora and genre, have come up with different, and sometimes contradictory, figures:

- It is estimated that we use 1.08 novel figures of speech and 4.08 idioms per minute, based on the analysis of approximately 200,000 words from political debates, taped psychotherapy sessions, and written adult and student compositions (Pollio *et al.*, 1977, pp.6-9). Cooper (1998, p.255) uses these figures to estimate that if people speak for about 4 hours/day, then they would use 4.32 novel figures of speech and 980 idioms.
- In three hours of recorded and transcribed conversational television data, idioms were used at the rate of about 3 per minute (Cooper, 1998, p.256).
- Three hours of television situation comedies and action-adventure shows in the United States yielded over 70 different idioms, without counting two-word verb constructions (Irujo, 1986b, p.241).
- Contrary to the belief of six language unit [primary school] teachers that they rarely used idioms in the classroom, this study reveals an average usage by these teachers of 1.73 idioms per minute (Kerbel and Grunwell, 1997, p.113).
- The whole corpus [trials, therapy talks, White House transcripts] had only 92 idioms which equals a ratio of one idiom per 1,150 words or as expressed in time, one idiom every four and a half minutes (Strassler, 1982, p.81).
- Moon (1998) compares the frequencies of 24 idioms in the Bank of English [over 400 million words of spoken and written English] with their frequencies in the 20 million word subcorpus of conversation, and states that “It appears that people may be impressionistically over-reporting high incidences of idioms in conversation” (pp.72-73).

Furthermore, the number of ‘idioms’ claimed to occur in written and spoken English is likely to depend, once again, on the definition and genre being used:

Although time does not allow a thorough search of figuratives and ONCEs in this study, a quick search of the BNC reveals that while some of them marked as frequent in *Cambridge* occurred more often than core idioms, others marked as frequent did not. Both are shown here in Table 53:

Table 53 BNC occurrences of ‘idioms’ marked as frequent in *Cambridge*

‘Idioms’ marked as frequent with many occurrences	‘Idioms’ marked as frequent with few occurrences
<i>on the other hand</i> – 5311 occurrences in 1613 texts	<i>a field day</i> – 58 occurrences in 55 texts
<i>in touch</i> – 1804 occurrences in 895 texts, <i>out of touch</i> – 196 occurrences in 166 texts	<i>straight face</i> – 53 occurrences in 43 texts
<i>come/comes/came/coming to terms with</i> – 834 occurrences in 587 texts	<i>work cut out</i> – 52 occurrences in 48 texts
<i>at the end of the day</i> – 760 occurrences in 483 texts	<i>on the off (-) chance</i> – 40 occurrences in 35 texts
<i>in the wake of</i> – 731 occurrences in 456 texts	<i>par for the course</i> – 38 occurrences in 35 texts
<i>make(s)/made/making the most of</i> – 569 occurrences in 410 texts	<i>off (-) the (-) cuff</i> – 35 occurrences in 29 texts
<i>in the long run</i> – 512 occurrences in 355 texts, <i>in the short run</i> – 122 occurrences in 66 texts	<i>snowed under</i> – 19 occurrences in 19 texts
<i>in the long term</i> – 436 occurrences in 330 texts, <i>in the short term</i> – 400 occurrences in 293 texts	<i>nail on the head</i> – 12 occurrences in 12 texts

A more comprehensive search of this and other corpora would give more exact frequency figures, but this quick search reveals that while some figuratives and ONCEs occur more often than core idioms, others do not. An examination of the ‘idioms’ marked as frequent in the two dictionaries which give a frequency indicator – *Cambridge* and Collins COBUILD – revealed that only 79 were shared by both dictionaries (see Appendix 8). Out of the 79, only 3 (*at loggerheads*, *take the mickey/mick*, *out and out*) would be judged to be core idioms, with the rest being figuratives and ONCEs. Both language teachers and learners would benefit by knowing which of the many idiomatic MWUs – core idioms, figuratives, and ONCEs – occur most frequently, and therefore merit their attention first.

Next we will look at how the results of the search for core idioms are listed on the attached CD, and in the appendices.

5.15 LISTING THE RESULTS IN APPENDICES

The decision was made to list all core idioms together, as in a dictionary. As explained, the overall results are given in the same format (see 5.8.1). However, as the corpus search results, showing the infrequency of core idioms, may be of interest to both language teachers and learners, the results and number of occurrences are listed in Table 29, as well as separately in different appendices according to frequency:

- Appendix 1: The 20 most frequent core idioms – 48-487 occurrences
- Appendix 2: Core idioms with 31-47 occurrences – 7
- Appendix 3: Core idioms with 21-30 occurrences – 4
- Appendix 4: Core idioms with 11-20 occurrences – 22
- Appendix 5: Core idioms with 10 or fewer occurrences – 34
- Appendix 6: Core idioms with zero occurrences – 17
- Appendix 7: The complete alphabetical list of core idioms and borderlines – 104

As can be seen, despite the limitations of basing frequency on a single corpus it is clear that core idioms do not occur frequently. However, it seemed useful to work out the ratio of each category. The average ratio was calculated by adding together the total number of occurrences of all the core idioms in each category, then dividing this number by 100 million (the number of words in the BNC), and finally multiplying this number by 100. Table 54 below gives the results, with the frequency of occurrence divided into different categories, from

the highest occurrence to the lowest, as well as both the number of core idioms and the average ratio of each category:

Table 54 Frequency of occurrence and number / ratio of core idiom occurrences

Appendix number and frequency of occurrence	Number of core idioms in each category	Number of occurrences in each category added together	Average ratio of occurrences in each category per 100 million words
Appendix 1: 48 or more occurrences	20	2345	.023
Appendix 2: 31 to 47	7	273	.003
Appendix 3: 21 to 30	4	92	.0009
Appendix 4: 11 to 20	22	319	.003
Appendix 5: 10 or fewer	34	154	.002
Appendix 6: 0 occurrences	17	0	0
TOTAL	104	3182	.032

The table shows that there are 2345 occurrences in the BNC corpus of the top 20 idioms. That means there are about 2 occurrences per million words or one occurrence of one of these 20 core idioms per 138 pages, if we assume that on average each page contains 300 running words. In total, core idioms occur at a density of 80 per 1 million running words or 1 per 1,250 running words, or 1 core idiom every 42 pages.

As mentioned, the search for core idioms in the BNC has shown that this particular category of idiomatic MWUs – core idioms – does not occur frequently enough to merit inclusion in any high frequency list. This has implications for their teaching and learning. The corpus search proves that rather than all idioms being frequent, this category which I have called core idioms, are in fact infrequent. In other words, it may not be worth either language teachers or learners spending time on them. Instead, learners can be taught vocabulary learning strategies such as guessing from context, as well as the effective use of a good idiom dictionary for the occasions when they may encounter one. More important for both language teachers and learners is research to establish a frequency list of the other two categories of idiomatic

MWUs – figuratives and ONCEs – many of which occur much more frequently than core idioms (see Table 53). Also useful might be suggestions for learning all three categories of idiomatic MWUs – starting with core idioms, should the need or desire exist.

5.16 APPLYING SECOND LANGUAGE ACQUISITION (SLA)

THEORY TO LEARNING IDIOMATIC MWUs

The theories of Second Language Acquisition (SLA), in particular as it relates to vocabulary learning, could therefore be applied to the learning of idiomatic MWUs.

To begin with, research shows that some idiomatic MWUs which have equivalents in the learners' L1 may be learnt easily, such as the 80% of expressions around 'to take' which have an idiomatic equivalent in French, 20% of these having a word-for-word equivalent. The same study shows that while learning is helped, acquisition can in fact be hindered if an idiomatic MWU is similar in the learners' L1 as it may pass unnoticed (Arnaud & Savignon, 1997).

Next, second language acquisition (SLA) theory supports the claim that a learner's knowledge can move from explicit (taught) to implicit (known/understood) if the learner has reached the necessary stage of acquisition, and is sufficiently motivated (Ellis, 1997). So first the learner has to be ready, and then perhaps to be taught, to 'notice' (Schmidt & Frota, 1986;

Schmidt, 1990; Ellis, 1994) the vocabulary items – in this case, the idiomatic MWU.

This ‘noticing’ is likely to first occur through reading. It has been shown that learning to read in a second language centrally involves learning words (Grabe & Stroller, 1997) and that extensive reading – including that of authentic materials – contributes to L2 vocabulary acquisition of both adults and children (Coady, 1997). Therefore, reading materials, preferably those containing authentic examples of idiomatic MWUs, could be used.

It has been argued that vocabulary, including core idioms, should not be taught in a ‘lexical set’, out of context (Nation, 2000). Yet according to McCarthy (1998), while idioms are highly interactive and best studied in context, they tend to be taken out of context and taught as separate items. Because of the difficulty of constructing the appropriate interactional climate for the teaching of idioms, McCarthy proposes the raising of students’ awareness of idiom usage as a first step. Students responded positively to an approach to learning idioms that began with consciousness-raising and moved to idioms used in authentic discourse contexts (Simpson & Mendis, 2003). Students also showed a preference for seeing how idiomatic MWUs were used in corpus examples (*ibid*).

In addition, certain ‘tricks’ or techniques can be used for vocabulary learning, including the learning of idiomatic MWUs. These include guessing in context, keyword technique, especially with words that can be perceived visually (Nation, 2001; Hulstijn, 1997), mnemonics (Nation, 1990, 2001; Hulstijn,

1997), communicative activities (Nation & Newton, 1997), plus exercises involving strong word partnerships and correct word partnerships, matching half phrases, putting idiomatic MWU parts into the correct order, completing sentences by adding a word missing from the idiomatic MWU, matching half idiomatic MWUs with the correct half and with their meanings and so on (Lewis, 1997).

Therefore, if the desire to learn some idiomatic MWUs exists, including even the infrequent core idioms, some guidelines for their comprehension, retention, and usage may be useful.

5.17 LEARNING CORE IDIOMS

It could be argued that this small number of low frequency core idioms (and of some figuratives/ONCEs, see Table 53) does not deserve teaching time. However, as frequency is not the only motivation for learning a specific vocabulary item, the desire to learn some core idioms may exist, so some suggestions related to their learning may be appreciated.

- 1) As a core idiom is both non-compositional and non-figurative, learners cannot apply techniques such as using word parts nor use a ‘figure’ to help them understand the MWU. They can, however, like with any vocabulary, use contextual clues to try to guess the meaning of the core idiom.
- 2) If this strategy isn’t successful, learners can be taught the effective use of a good idiom dictionary. As shown (see 5.7.1), idiom dictionaries do not list entries the same so learners may need to be taught how to find

them in a variety of paper dictionaries, especially if their electronic dictionaries do not include the particular core idiom or do not provide a satisfactory explanation.

3) As suggested, learners can be taught to ‘notice’ or pay attention to the core idiom, including time spent teaching them how to:

- identify the core idiom when they see it written or hear it spoken – in other words, notice where the MWU begins and ends, and how many words it contains
- locate the core idiom in the idiom dictionary – by the first word, by the noun (herring) or adjective (red), by the ‘head word’ written in bold in the index, and so on
- learn its grammatical behaviour, especially using corpus (BNC) examples – if it begins with a verb, whether the verb can be substituted with a similar verb (*make / create sth out of whole cloth*), whether the tense can change (*make/makes/made/without making any bones about it*), whether it can occur in both active (*sell sb a bill of goods*) and passive voice (*get sold a bill of goods*), whether it can be used as a noun (*is the bee’s knees*) or an adjective (*bees’ knees service*), and so on
- recognise its form and whether this form is frozen or allows variations, again using corpus (BNC) examples – in spelling (*bang / banged to rights*), hyphenation (*odds and sods, odds-and-sods*), apostrophes (*the cat’s pyjamas, the cats’ pyjamas*), singulars/plurals (*eat your heart / hearts out*), internal (*cut a particular dash*) or external modification (*absolutely all in*), truncation (*Butter wouldn’t melt in sb’s mouth – Butter wouldn’t melt*) and so on
- notice its collocations in corpus (BNC) examples – whether *like* is a clue that some sort of comparison follows, including core idioms such as *like the clappers* and *like nobody’s business*, and whether verbs such as ‘dismiss’ and ‘reject’ are likely to be followed by core idioms such as *out of hand*, and so on

4) Also helpful to learners might be an understanding of the historical/etymological background or origin of core idioms and figuratives (see, for example, Ammer, 1992, 1997; Cryer, 2001; Donald, 1994; Flavell

& Flavell, 1992, 1993; Speake, 1999, *Why Do We Say It*, 1985) as understanding how the MWU came into existence might serve as a 'hook' to help learners remember it better. It might even provide a 'figure' and move the MWU from the core idioms category into the figuratives one. Despite occurring infrequently, the core idiom may be crucial to the understanding of a passage or the tone of a message (in a newspaper article, a story, a conversation, a lecture, and so on) because of how it is used, where it is used, or the way in which it is used (see Arnaud & Savignon, 1997).

5) As core idioms are also collocations, language teachers may want to investigate the teaching of collocations (see Lewis, 2000) or apply some of Nation's (2001, pp.329-332) scales which indicate what is involved in learning collocations:

- frequency of co-occurrence (how frequently do 'nip and' combine with 'tuck')
- adjacency (how often does 'knees' follow 'bee's')
- grammatically connected (if 'splice' and 'mainbrace' are found in the same sentence, do they always co-occur as this MWU)
- grammatically structured ('at sixes' might be 'loosely related' but 'at sixes and sevens' would be 'well-structured')
- grammatical uniqueness (some structures such as 'come the raw prawn' are 'grammatically unique' rather than 'grammatically regular')
- grammatical fossilisation ('by and large' does not allow any change to its form)
- collocational specialisation (how specialised are 'trip the' and 'light fantastic' compared to their co-occurrence with other words)
- lexical fossilisation (while 'the cat's' can combine with 'whiskers', 'pyjamas' and 'miaow', not to mention non-idiomatic occurrences, 'like nobody's' is more likely to be fossilised with 'business')
- semantic opaqueness (how semantically opaque is 'red herring' compared to 'red dress')

- uniqueness of meaning (while ‘cut a dash’ and ‘eat your heart out’ have only the idiomatic meaning each, ‘kick the bucket’ and ‘off the wall’ have both the idiomatic and literal meaning)

Learners can therefore treat core idioms like any other single word, and despite their infrequent occurrence, may still wish to learn them. As noted some considerable time ago:

A student may learn grammar and acquire sufficient vocabulary to communicate, but without a working knowledge of such expressions...his spoken and written English will remain stilted and foreign-sounding. His reading will be appreciably slower and his comprehension will suffer. For these specified reasons the teaching of idioms and figures of speech should not be neglected, but should be made a part of the study of grammar and vocabulary for those students who learn English as a second language.

(Adkins, 1968, p.149)

Language teachers must continue to respond to the needs and wishes of learners with regard to the learning of vocabulary, including the learning of all three categories of idiomatic MWUs. This is particularly true of the learning of some figuratives and ONCEs.

5.18 LEARNING FIGURATIVES AND ONCEs

With regard to the learning of figuratives and ONCEs, more techniques are available to language learners. Moon (1997, p.58), however, comments that because many metaphorical MWUs are “marked, infrequent, and generally considered ‘difficult’, they may be taught sparingly as receptive vocabulary items”. Rather than sideline the teaching of figuratives and ONCEs to a Friday afternoon ‘fun’ activity, they could be incorporated into the target vocabulary

of every programme, starting with the more frequent ones. As Simpson & Mendis (2003, p.432) note:

The discovery of a significant number of idioms in a corpus of academic speech and, more importantly, the evidence that they perform a variety of important pragmatic functions provides the rationale for including them in an EAP curriculum.

And rather than being ‘difficult’ to learn, figuratives are more fun and should be easier to learn because they are semantically motivated and can often be ‘hooked’ to discourse functions (see McCarthy, 1998, pp.145-149; Simpson & Mendis, 2003, pp.427-432). In addition, as with core idioms, understanding even an infrequent figurative or ONCE may be crucial to understanding the main message of what is being said or written. A quick look at the range of figuratives and ONCEs used on television (see 4.9.6, 4.9.7) lends support to this claim.

As well as considering what is involved in learning vocabulary, including core idioms, language teachers can try a number of different strategies, again based on SLA and vocabulary learning theory, to make their learners more aware of both conceptual and linguistic metaphors (figuratives and ONCEs) and figurative language in general.

Some of these strategies follow:

- 1) Help learners gain greater independence by making them aware not just of metaphors but also of the metaphoric base of much of our language (Lakoff & Johnson, 1980). Even beyond our language, it has been argued that metaphors are “a matter of thought not language” (Lakoff & Turner, 1989, p.107) and are a means by which people make sense of the world around them. Perhaps because we do this automatically and unconsciously,

language teachers do not always recognise its importance in the language classroom. For more than a decade, both linguists and language teachers have stressed the need to teach ESL/EFL learners to work with metaphor in the target language (Cooper, 1999; Hatch & Brown, 1995; Lennon, 1998; Low, 1988; MacLennan, 1993, 1994). There is, however, little evidence that this has filtered down to the language classroom, perhaps because MWUs such as figuratives and ONCEs are regarded as linguistic expressions independent of any conceptual system and are isolated from each other at the conceptual level (Boers, 2000a, 2000b, 2001; Boers & Demecheleer, 2001; Gibbs, 1997; Kovecses & Szabo, 1996), or perhaps because they are considered by teachers simply too difficult for learners. However, research has shown that “complexity level in L2 metaphor interpretation appears to be quite independent from proficiency level in L2” (Johnson, 1996, p.227) and research into L1 showing the metaphoric base of so much of our language “has already produced results which L2 professionals cannot afford to ignore” (MacLennan, 1994, p.108). For example, using the ‘mind is a container’ metaphor, students can relate to having thoughts ‘in mind’, ‘on your mind’, ‘pass through your mind’, ‘called to mind’, and ‘put out of your mind’ (MacLennan, 1994, p.105). Giving learners practice in identifying the conceptual metaphor behind a variety of linguistic expressions will help them gain the skills to do this independently as well.

- 2) Give learners a greater understanding of figurative language and metaphors, and introduce them to language teaching textbooks aimed at this (see, for example, Lazar, 2003; Wright, 1999). Teachers could start

with the more familiar similes and proverbs, and look at similar MWUs used in the learners' first language.

- 3) Help learners become aware that many English proverbs have become so well-known, institutionalised and predictable that they are frequently truncated as Table 55 shows with examples from the BNC:

Table 55 Examples of truncated proverbs found in the BNC

Truncated proverbs which may not be understood by ESL/EFL learners			
When in Rome	Hell hath no fury	The grass is (always) greener	An eye for an eye
If the cap fits	When the cat's away	People (who live) in glass houses	A bird in the hand
A bad penny	You scratch <i>my</i> back	Don't count your chickens	The silver lining
Two's company	You've made your bed	(Don't)shut the stable door	(the/Mr)new broom

- 4) Help learners begin to recognise figuratives and ONCEs used in the media, often in their truncated form. For example, truncated proverbs such as 'new broom' can also be used visually, as demonstrated in this cartoon from a Wellington daily newspaper, *The Dominion* (11 December, 2001). For learners, Figure 1 shows that understanding the cartoon and appreciating the humour involves not only understanding the full proverb, but also knowing that the proverb can be truncated, can be visual as well as verbal, and can be implicit as well as explicit, not to mention understanding the political implications and the people (newly appointed Defence Force Chief Ferguson) involved in the story. Language teachers can help learners become aware of all of this.

Figure 1 Example of a ‘visual idiom’ of *new broom*



5) In addition, help learners become aware that the adjective ‘proverbial’ has two meanings:

- something that is ‘connected with or resembling a proverb’ (65% of BNC occurrences)
- something that is ‘characteristic or notorious’ (35% of BNC occurrences)

From similes and proverbs, learners can move to other figuratives and ONCEs, and can be given strategies to make sense of them.

6) Help learners use the etymological information as a strategy for learning figuratives (see Boers & Eyckmans, in press) and ONCEs. Knowing how the figurative came to be used may help learners both understand and remember it by giving them an ‘image’ to use. In addition, when learners realise that the meaning of many figuratives is based on a ‘conceptual’ metaphor (LIFE IS A JOURNEY), they begin to understand the linguistic metaphors/figuratives that result (*moved into the fast lane, take a back seat, gone off the rails*, and so on) which become “imageable” (Lakoff, 1987). They can also look for similar conceptual metaphors in their L1.

7) Help learners become familiar with the various themes from their target culture which idioms are related to, such as sports (*knock sb for six, go to bat for someone, be par for the course*), gambling (*play your cards close to your chest, follow suit*), sailing (*go overboard, be plain sailing, be in the doldrums*), and so on. Once again, this can be compared to themes in the learners' L1.

8) As well as providing the necessary cultural information, help learners to recognise the 'untruth' and look for the 'figure' in figuratives so they can reinterpret the MWU in the known circumstances, such as these examples show:

- give sth/sb the green light – learners know that a green light in traffic gives permission to proceed, so if a project or plan is 'given the green light', they can recognise the 'untruth' with the 'green light' in relation to their request, and reinterpret it in the circumstances they are in to mean 'proceed with their request'
- add fuel to the fire – learners know what happens to a fire if any sort of fuel is added to it, so they can recognise the 'untruth' with fuel and fire in relation to what is being said, and reinterpret it in the circumstances to mean 'make the situation bigger/faster', something which in the circumstances is likely to be negative

This is not as straightforward as these examples may indicate. Learners can be shown the cognitive style differences involved. As mentioned, many figuratives are imageable, and research has shown that using "imagers" may be more effective than using "verbalisers" to associate a mental image with an figurative (Boers & Littlemore, 2000). The learners' ability to use the underlying conceptual metaphor to make sense of the resulting figurative may depend on variables such as understanding the

individual words, including polysemy, as well as knowing whether the same conceptual metaphor exists in their L1 (Charteris-Black, 2002).

9) Help learners see how all idiomatic MWUs are used in ‘real language’ by providing corpus-based pedagogical materials, so learners can see not only the range of genres in which they occur, but also the pragmatic functions they perform (Simpson & Mendis, 2003).

- ‘on that note’ and ‘shift gears’ are typically used at a discourse boundary, signalling the end of one topic or episode and a transition to another
- ‘lost my train of thought’ is almost always used in a situation of disfluency

10) Help learners get practice with figuratives that have a less obvious ‘figure’ to reinterpret, such as the following:

- (be) a dirty old man – learners may guess that if man is not actually physically ‘dirty’, he must be ‘dirty’ in some other way, and using the circumstances can be helped to reinterpret or guess in what way(s) the man is thought to be ‘dirty’
- throw a wobbler/wobbly – learners may guess that if a ‘wobbler/wobbly’ does not exist as a noun, and if a person does not actually throw anything, that they have to reinterpret the whole MWU in the known circumstances to guess its intended meaning

If the minimum context does not help learners to reinterpret the figurative, they may need to be given a little more information/context until they are able to successfully guess the correct meaning.

11) Help learners get practice with ONCEs, so they know that only one word is idiomatic, and can practise guessing the meaning from the circumstances, including examples like the following:

- stick to your knitting – if learners know that the first three words have their literal meaning, they can once again begin to guess what ‘knitting’ means in the context in which it is being used
- filthy rich – if learners know that the final word has its literal meaning, they can once again begin to guess what ‘filthy’ means in the context in which it is being used
- give sb a buzz / get a buzz from – if learners know that only one word of the MWU is idiomatic, they can begin to guess the different meanings of ‘buzz’ in the context in which they are being used

With practice, learners can gain some independent learning strategies and begin to do this outside of the classroom.

12) Help learners become aware that language – especially figurative language – can be used to make us feel positive or negative towards a subject. Carter, Goddard, Reah, Sanger & Bowring (1997, p.87) show how the “metaphoric use of language reflects our cultural attitudes” with these examples:

- Mutton dressed as lamb / She’s a bitch / She’s a cow / She’s a slag
- America is our ally / The rape of Sarajevo / Cuba is the victim of the American blockade / The evil empire
- Surgical strike / Ethnic cleansing / Friendly fire (killed by your own side)
- The economy is ailing / The trade figures are healthy this month / The government does not realise how much is at stake

Media coverage of the recent 2003 war in Iraq – softened to ‘conflict in Iraq’ by Television New Zealand – has more examples of this. Learners would benefit by discussing how such cultural attitudes are used in their L1 compared to their target language.

These are only some strategies language teachers can use. When learners are made aware of figuratives and ONCEs, similes and proverbs – both full and truncated, the role of metaphors and metonymy, the conceptual metaphors on which much of our language and thought is based, and how language shapes our attitudes, they can begin to have a better understanding of the idiomatic MWUs used around them. All of these strategies are aimed at making our ESL/EFL learners more aware of the role of figurative language and more independent as language learners operating in their target culture.

In addition, as well as focussing on strategies for learning figuratives and ONCEs, other suggestions for the teaching of collocations (see, for example, Aisenstadt, 1981; Bahns & Eldaw, 1993; Farghal & Obiedat, 1995; Lewis, 1993, 2000; Sinclair, 1991a; Wray, 2000) can be examined to see which ones also apply to the teaching of figuratives and ONCEs. Furthermore, suggestions for the teaching of vocabulary in general, and ‘idioms’ in particular, can be studied to see if they can be applied to the teaching of figuratives and ONCEs (see, for example, Auld, 2001; Burns & de Silva Joyce, 2001; Celce-Murcia & Rosenweig, 1979; Cooper, 1998; Cowie, 1992; Hatch & Brown, 1995; Holmes & Moulton, 1993; Honeyfield, 1977; Irujo, 1986a, 1986b, 1993; McCarthy, 1984, 1991; McCarthy & O’Dell, 1994, 1999, 2002; Meara, 1980; Moon, 1997; Nation, 1990, 2001; Naylor, 2002a, 2002b; Shaptozhvili, 2002; Sokman, 1992; Waring, 1997; Watson, 1998; Wright, 1999).

Having looked at some of the results of the corpus search and at some suggestions for teaching idiomatic MWUs, we will now conclude the study,

discuss implications for further research, and summarise what the study has attempted to accomplish in Chapter 6.

CHAPTER 6

SUMMARY OF CONCLUSIONS AND IMPLICATIONS FOR FURTHER RESEARCH

6.1 SUMMARY

The primary intent of this study was to find, or establish, a clear definition of an ‘idiom’ that would distinguish it from a ‘non-idiom’. I began the study by surveying the literature to try to find this definition but not having found one that made a clear distinction, an attempt was made to create one. The result was a new definition of a ‘core idiom’ – an MWU (made up of at least two words which occur elsewhere) that was both:

1. non-compositional
2. non-figurative

Following this, I next developed a test to divide the present large collection of idioms into the three new categories: figuratives, ONCEs, and core idioms. I then applied this test to idioms from the same page of two different idiom dictionaries, where it proved to have a high degree of success at grouping these idioms into the three newly-created categories. This completed the first stage of the study.

Having set aside the first two and the most frequent categories – figuratives and ONCEs – I then selected the third category, core idioms, for the corpus search. This category was chosen because although the smallest category, these were the ones most likely to cause difficulty for ESL/EFL learners. The second stage of the study involved developing a comprehensive list of core

idioms to be used in the corpus search. The compilation of this list comprised a number of steps, starting with a study of ten different idiom dictionaries to choose a primary source dictionary. I chose the *Cambridge International Dictionary of Idioms* (McCarthy & Walter, 1998) as my primary source of core idioms because of factors such as the large number of entries claimed ('around 7,000'), the inclusion of some source information such as 'British, American, and Australian' justifying its claim to be 'international', the recency of publication (1998), and the inclusion of a frequency indicator.

After choosing the primary source, I next made a detailed and thorough search of seven of the remaining nine dictionaries and I perused the remaining two to see which additional core idioms could be added to the original list of core idioms from *Cambridge*. It turned out that, as might have been expected, dictionaries proved to be the best source of core idioms. In order to widen the search, I began to look at other sources of core idioms, starting with sources considered relevant to my students – both ESL/EFL teaching course books and vocabulary teaching books. The search was then widened to include core idioms mentioned in journal articles and book chapters written about them, and as well as two written sources known to be 'idiom'-rich, horoscopes and sports reports. Lastly, in order to find sources of 'spoken' English, I added two additional sources, core idioms from both scripted and unscripted television. I then searched all of these sources to see how many additional core idioms could be added to the original list in an attempt to produce the most comprehensive list of core idioms possible. Finally, in order to establish a frequency count and to gather a range of other information about each core idiom, I carried out a corpus search on the British National Corpus of all of

those MWUs which qualified as core idioms. The BNC was selected ahead of sources of New Zealand English because of its much larger size (100 million words), thus increasing the chances of finding sufficient occurrences of the low frequency target core idioms.

The corpus search resulted in a list of the most frequent core idioms, with the most frequent 20 core idioms having between forty-eight and four hundred and eighty seven occurrences each in the 100 million corpus, and comprising core idioms used as discourse markers (*by and large*), as vague language (*so and so, such and such, and what have you*), as nouns (*The Big Apple, red herring(s)*), and verbs (*take the piss/mickey, pull someone's leg*), adjectives (*at loggerheads, out and out*) and adverbs (*like the clappers, to a tee*). Core idioms are divided into frequency lists, ranging from 487 to zero occurrences, in Appendices 1 to 6 with the complete list of core idioms given alphabetically in Appendix 7. The results of the corpus search, including details about each one as well as examples of each, are given on the accompanying CD.

Results of the corpus search show that of the 104 core idioms and borderlines:

- none occurs frequently enough to merit inclusion on any high frequency word list
- 17 have zero occurrences in the BNC
- 40 have 0 occurrences in spoken English in the BNC
- 34 have fewer than 10 occurrences in the BNC
- 91 have no literal occurrences in the BNC
- 13 have literal occurrences in the BNC
- only 3 of the 13 with literal occurrences have more literal than idiomatic occurrences
- 20 occur at least twice as commonly in spoken than in written English in the BNC
- another 5 occur more commonly in spoken than in written English in the BNC

- 55 occur more commonly in written English in the BNC, including 22 which occur considerably more commonly
- 8 core idioms occur approximately equally commonly in both spoken and written English in the BNC
- the remaining ones occur more commonly in written English in the BNC

A corpus search of the much smaller Wellington Corpora (written and spoken) reveals that:

- *by and large* occurs frequently in both the BNC and the WC/WSC
- *so and so, serve sb right, take sb to task, red herring(s), touch and go* would occur about equally frequently if both corpora were the same size
- *such and such, take the piss, out and out, pout your foot in your mouth, have an axe to grind* were comparatively more frequent in the larger BNC
- *out of hand, pull sb's leg, (be) a piece of cake, come a cropper, make no bones about it* were comparatively more frequent in the smaller WC/WSC

With regard to the teaching of core idioms, an argument could be made that their infrequent occurrence established by a corpus search of the BNC indicates that language learners should not spend time on them. Instead, learners should develop skills in using a good idiom dictionary for the few occasions when they might encounter one. Another argument, however, is that at the higher levels, frequency becomes a more tenuous concept and that the “instinct of an experienced teacher” as to what is a useful MWU might be more valuable (Cornell, 1999, p.5). This is supported by the search of the BNC which showed that some core idioms – such as those listed in Table 56 below – were not found to be frequent, yet may be worthwhile teaching, depending on the age and needs of the learners:

Table 56 Examples of BNC core idioms numbering ‘10 or fewer’ and ‘20 or fewer’

Core idioms numbering 10 or fewer	Core idioms numbering 20 or fewer
<i>a fine kettle of fish</i>	<i>Bob’s your uncle</i>
<i>(be) the bee’s knees</i>	<i>kick the bucket</i>
<i>to boot</i>	<i>not cut the mustard</i>
<i>have a bone to pick with sb</i>	<i>to a T / tee</i>
<i>like nobody’s business</i>	<i>like the clappers</i>
<i>(be) nip and tuck</i>	<i>cold turkey</i>
<i>do your nut</i>	<i>odds and sods</i>
<i>(be) at sixes and sevens</i>	<i>Beat it!</i>
<i>cook sb’s goose</i>	<i>(be) all in</i>
<i>(be) off your face</i>	<i>Butter wouldn’t melt in sb’s mouth</i>
<i>sweet Fanny Adams / FA</i>	<i>chew the fat</i>

With regard to the teaching of the other two categories of idiomatic MWUs – figuratives and ONCEs – previous suggestions have been proposed for the selection of which ones to teach, including need, usefulness, productivity, currency and frequency, and ease (Yorio, 1980). Other suggestions based on SLA theory include starting with equivalent idiomatic MWUs in learners’ L1, getting learners to ‘notice’ various things about the MWU, providing authentic examples of the MWU used in context, such as reading materials, using techniques such as guessing in context, keyword and mnemonic techniques, using communicative activities and a variety of exercises including word partnership, matching, word order, adding missing words, and so on. A strong argument exists for the teaching of high frequency ones first, indicating the need for a corpus search to establish the frequency of these two groups. A brief look at any idiom dictionary would reveal that the majority of MWUs are in fact figuratives. We have already discussed strategies for learning these (see 5.17). The fact that there is some degree of systematicity is important. In other words, it is worth making learners aware of figurative language in general and getting them used to recognising the ‘untruth’ in the MWU and learning how to ‘reinterpret’ it in the circumstances in which it occurs.

Further, it has been argued (Laufer, 1992) that lexicographer's examples are often better for ESL/EFL learners than corpus examples. While this may be true for lower-level learners, a preference for corpus examples showing the MWU used in context has also been shown (Simpson & Mendis, 2003). The sort of language used on television (see 4.9.6, 4.9.7) and in the corpus indicates that both textbook writers and language teachers should be encouraged to balance teacher-made examples with this 'real life' language so that learners can successfully move from the classroom to the real world in terms of their language development.

In addition, while core idioms, figuratives, and ONCEs may have a limited role in certain types of text, such as scientific texts, and a much larger role in soap operas and horoscopes, it has been suggested that somewhere in between these two might be an area such as business journalism (Cornell, 1999) and academic discourse (Simpson & Mendis, 2003). It seems clear that need (for understanding, communicating with and being accepted by peers, for understanding academic texts, for understanding tertiary lectures, for reading the newspaper, for watching television dramas, and so on) may take priority for the learners in terms of what vocabulary they choose to learn. Clearly, the instinct of an experienced language teacher has a role to play in deciding which idiomatic MWUs learners should focus on, and helping them find ways to best learn them.

The important issue is this – however and whenever language teachers choose to teach core idioms, figuratives, and ONCEs, the awareness of them should be an integral part of vocabulary learning at all levels, something which should

help learners overcome their ‘idiomphobia’ (Irujo, 1986a, p.300) and begin to communicate naturally in their target language and culture.

Despite the work done previously on the teaching and learning of idioms, and the work in this study, there remain many areas of investigation for the future. I now briefly highlight some of these areas and suggest several avenues for further research.

6.2 IMPLICATIONS

While it is hoped that this study has contributed to the understanding of idiomatic MWUs, there remain many shortcomings which could be addressed by future research, and these along with the final summary, form the concluding section of this chapter.

Future research could focus on the following:

- 1) A similar corpus search could be done on the other two categories of figuratives and ONCEs to establish which ones are most frequent in order to know which ones to teach/learn first
- 2) A similar corpus search could be done using a variety of corpora such as targeted academic texts and lectures (see Simpson & Mendis, 2003), popular journalism, television dramas and documentaries, and so on, to determine which MWUs in all three categories are both frequent and useful across a variety of genres – in other words, their range – and which should be taught/learnt first

3) A corpus search could be done to determine the following:

- which age groups are using core idioms, figuratives and ONCEs most
- which particular MWUs is each age group using most
- how are these MWUs being used by different age groups in different genres (conversation, academic lectures, fictional texts, factual texts, newspapers, television news, and so on)
- which of the variety of forms of core idioms, figuratives and ONCEs are being used most frequently
- what patterns, collocations and associations of core idioms, figuratives, and ONCEs are the most common
- how language teachers and learners can take advantage of this information

4) A corpus search could be done to verify what other researchers, prior to corpus studies, have claimed about idioms:

- ‘idioms’ are used when making complaints (Drew & Holt, 1988)
- ‘idioms’ are used in topic transition (Drew & Holt, 1995, 1998)
- ‘idioms’ play a role in negotiating workplace encounters (Koester, 2000)
- ‘idioms’ are used as evaluative devices (McCarthy, 1998; Simpson & Mendis, 2003; Strassler, 1982)
- ‘idioms’ are used as a paraphrasing or observation-plus-comment device (McCarthy, 1998; Simpson & Mendis, 2003)

- ‘idioms’ are used to emphasise content or reinforce an explanation (Simpson & Mendis, 2003)
- ‘idioms’ are used to establish a sense of solidarity within a group of speakers (Simpson & Mendis, 2003)
- ‘idioms’ are used by the speaker to reduce the formality and transactional nature of academic discourse (Moon, 1998a, Simpson & Mendis, 2003)
- ‘idioms’ are used to establish social hierarchy or to test a hearer’s opinion (Strassler, 1982)
- ‘idioms’ are used as descriptive devices, to highlight contrast or for economy of expression (Simpson & Mendis, 2003)
- ‘idioms’ are used as metalanguage, to organise texts by signalling connections between them (Fernando, 1996; Moon, 1998a; Simpson & Mendis, 2003)
- ‘idioms’ are used as ‘cross-functioning’ devices, serving more than one function (Moon, 1998a; Simpson & Mendis, 2003)
- how this knowledge benefits both language teachers and learners

5) Classroom research could be done to determine whether teaching learners how to ‘recognise the untruth’ in the figurative/ONCE and ‘reinterpret it in known circumstances’ helps them learn to correctly guess the meaning of the MWU

- 6) Classroom research could be done to build on what Boers & Eyckmans (in press) found, that a cognitive semantic approach (etymological elaboration) to teaching figuratives is beneficial to ESL/EFL learners
- 7) Classroom research could be done to determine whether teaching learners the conceptual metaphors which underlie so much of our language (Lakoff & Johnson, 1980) helps them learn to interpret and correctly guess the meaning of idiomatic MWUs (see, for example, Kovecses & Szabo, 1996; Boers, 2000)
- 8) Classroom research could be done to determine how ESL/EFL students prefer to learn core idioms, figuratives and ONCEs, as well as how they are most successful at learning them (see, for example, Boers & Eyckmans, in press)
- 9) Classroom research could be done to determine the sorts of errors learners are making when using core idioms, figuratives, and ONCEs in order to know how to help them avoid, or correct, these errors (see, for example, Cornell, 1999)
- 10) Research could be done on how idiomatic MWUs are used in other languages, compared to how they are used in English, to find similarities and differences (see, for example, Boers & Demecheleer, 2001; Charteris-Black, 2002)

- 11) Research could be done on where and how often core idioms, figuratives, and ONCEs are used in a variety of situations, in preference to their non-idiomatic (literal) alternatives so language teachers/learners know which to teach/learn first
- 12) Research could be done on whether core idioms, figuratives, and ONCEs share certain grammatical patterns and associations and whether knowledge of this aids ESL/EFL students in the learning of them
- 13) Research could be done on whether core idioms, figuratives, and ONCEs share certain collocational patterns, and whether knowledge of this aids ESL/EFL students in learning them, or in predicting their use
- 14) Research could be done on how many figuratives and ONCEs occur in only a fixed or frozen form, how many allow variations, and which variations are used most often so language teachers/learners can take advantage of this information
- 15) A corpus search could be done on specifically New Zealand data, both written – academic textbooks used in different high school and university courses, articles in popular magazines and New Zealand daily and weekly newspapers – and spoken sources – academic lectures, New Zealand television and radio news and so on – to determine the same things as listed above

16) A corpus search of newspapers articles/journalism could be done to verify how often verb-noun/adjective-noun/adverb-verb (restricted) collocations are used and which ones are used most often, compared to how often core idioms, figuratives and ONCEs are used (Cowie, 1992) so language teachers and learners can take advantage of this information

17) A corpus search could be done of the frequently-used – and also difficult for ESL/EFL students to learn – category of phrasal verbs, both idiomatic and non-idiomatic, so this information can contribute to their teaching/learning

18) A corpus search could be done of the other categories which do not meet the core idiom criteria, including compound word MWUs, foreign word MWUs, fictional character MWUs, literal MWUs, literary MWUs, and collocational fixed phrase MWUs so language teachers and learners can judge their importance and role in vocabulary learning

These are only a few of the research projects that would make a useful contribution to the field of second language acquisition as related to the learning of idiomatic MWUs. A survey of both classroom teachers and learners is likely to come up with several others related to the field. The more we are able to find out about core idioms, figuratives, and ONCEs – their frequency and use in different genres in particular – the more we can assist both language teachers and learners to deal with these MWUs, and the greater

will be our contribution to the field of vocabulary learning in second language acquisition.

With regard to the research questions posed in the Introduction, I have attempted to develop clear criteria which can be applied to idioms to divide them into three distinct groups. I have established the frequency of one of these groups, that of *core idioms*. And I have suggested how the theory of SLA could be applied to the teaching, learning and acquisition of idiomatic MWUs.

While the findings from this study have permitted various insights into the field of idiomatic MWUs, it is my hope that others will take up the challenge of carrying out more research in this area in the future, to further advance our knowledge of the teaching, learning and acquisition of these MWUs, in order to further contribute to the field of second language acquisition.

APPENDICES

APPENDIX 1: The 20 most frequent core idioms

(occurring 48-487 times)

1. <i>by and large</i> – 487	11. <i>out (-) and (-) out</i> – 72
2. <i>so (-) and (-) so</i> – 327	12. <i>take the mickey</i> – 71
3. <i>such(-)and(-)such</i> – 196	13. <i>(be) at loggerheads</i> – 63
4. <i>out (-) of (-) hand</i> – 141	14. <i>pull someone's leg</i> – 60
5. <i>take the piss</i> – 137	15. <i>touch (-) and (-) go</i> – 53
6. <i>and what have you</i> – 136	16. <i>the Big Apple</i> – 52
7. <i>serve sb right</i> – 101	17. <i>cut no ice with sb</i> – 50
8. <i>take someone to task</i> – 92	18. <i>come a cropper(s)</i> – 49 [<i>come a purler</i> – 1]
9. <i>red herring(s)</i> – 87	19. <i>(be) a piece of cake / piss</i> – 50 [<i>(be) a piece of cake</i> – 43 / <i>piss</i> – 7]
10. <i>(be) beside yourself</i> – 72	20. <i>put your foot in it / in your mouth</i> – 48 [<i>in it</i> – 37, <i>in your mouth</i> – 11]

APPENDIX 2: Core idioms occurring 31-47 times

1. <i>(have) an axe to grind</i> – 47	5. <i>Uncle Sam</i> – 35
2. <i>make no bones about it</i> – 44	6. <i>and what not</i> – 33
3. <i>white elephant(s)</i> – 43	7. <i>(go) by the board(s)</i> – 32
4. <i>(be) all of a piece</i> – 38	

APPENDIX 3: Core idioms occurring 21-30 times

- | | | | |
|-------------------------------------|----|---|----|
| 1. <i>So long!</i> – | 24 | 3. <i>not hold a candle to sb/sth</i> – | 23 |
| 2. <i>(sb) eat your heart out</i> – | 23 | 4. <i>(be) off the wall</i> – | 22 |

APPENDIX 4: Core idioms occurring 11-20 times

- | | | | |
|--------------------------------------|----|---|----|
| 1. <i>pull the other one/leg</i> – | 19 | 12. <i>cold turkey</i> – | 14 |
| 2. <i>trip the light fantastic</i> – | 18 | 13. <i>kick the bucket</i> – | 13 |
| 3. <i>cut a dash</i> – | 18 | 14. <i>the cat's whiskers</i> – | 13 |
| | | <i>(the cat's pyjamas</i> – | 2) |
| 4. <i>to a T/tee</i> – | 18 | 15. <i>give sb the bird</i> – | 13 |
| 5. <i>by and by</i> – | 17 | 16. <i>(be) all in</i> – | 13 |
| 6. <i>push the boat out</i> – | 16 | 17. <i>a French letter</i> – | 12 |
| 7. <i>gird (up) your loins</i> – | 16 | 18. <i>not cut the mustard</i> – | 12 |
| 8. <i>(go) hell for leather</i> – | 15 | 19. <i>odds and sods</i> – | 12 |
| 9. <i>cock(-)a(-)hoop</i> – | 15 | 20. <i>like the clappers</i> – | 12 |
| 10. <i>Bob's your uncle</i> – | 14 | 21. <i>Beat it!</i> – | 12 |
| 11. <i>chew the fat/rag</i> – | 14 | 22. <i>butter wouldn't melt in sb's mouth</i> – | 11 |

APPENDIX 5: Core idioms occurring 10 or fewer times

1. <i>sweet Fanny Adams/FA</i> – 10	18. <i>shoot the breeze</i> – 4
2. <i>to boot</i> – 10	19. <i>heap coals of fire on sb's head</i> – 3
3. <i>knock into a cocked hat</i> – 9	20. <i>swing the lead</i> – 3
4. <i>cook sb's goose</i> – 8	21. <i>(have) bats in the belfry</i> – 3
5. <i>do your nut/bun/prunes</i> – 8	22. <i>sell sb a bill of goods</i> – 2
6. <i>Old Glory</i> – 8	23. <i>splice the mainbrace</i> – 2
7. <i>cut up rough</i> – 8	24. <i>a fine/pretty kettle of fish</i> – 2
8. <i>pop your clogs</i> – 8	25. <i>set sb/sth by the ears</i> – 2
9. <i>bang to rights</i> – 8	26. <i>take the rise out of sb</i> – 2
<i>dead to rights</i> – 1	
10. <i>(be) at sixes and sevens</i> – 7	27. <i>lay an egg</i> – 1
11. <i>(be) the bee's knees</i> – 7	28. <i>all ends up</i> – 1
12. <i>like nobody's business</i> – 7	29. <i>how's your father</i> – 1
13. <i>cut a/the rug</i> – 6	30. <i>make sth out of whole cloth</i> – 1
14. <i>(be) off your face</i> – 5	31. <i>all gas and gaiters</i> – 1
15. <i>get your rocks off</i> – 5	32. <i>make your number</i> – 1
16. <i>have a bone to pick</i> – 4	33. <i>half seas over</i> – 1
17. <i>(be) nip and tuck</i> – 4	34. <i>hop the twig/stick</i> – 1

APPENDIX 6: Core idioms with zero occurrences in the BNC

1. <i>Box and Cox</i>	7. <i>the ghost walks</i>	13. <i>part brass rags with</i>
2. <i>bust sb's chops</i>	8. <i>go bung</i>	14. <i>row crosshanded</i>
3. <i>clean sb's clock</i>	9. <i>the bitch goddess</i>	15. <i>tie one on</i>
4. <i>eat crow</i>	10. <i>lay out in lavender</i>	16. <i>over the transom</i>
5. <i>to a fare-thee-well</i>	11. <i>peg it</i>	17. <i>wear the green willow</i>
6. <i>buy the farm</i>	12. <i>come the raw prawn</i>	

APPENDIX 7: Complete alphabetical list of core idioms and borderlines

(104) (bold indicates zero occurrences)

<i>(be) all in</i>	<i>lay an egg (b/f)</i>	<i>make your number (b/f)</i>
<i>(have) an axe to grind (b/f)</i>	<i>a white elephant (b/f)</i>	<i>do your nut / bun / prunes</i>
<i>(have) bats in the belfry (b/f)</i>	<i>all ends up (b/f)</i>	<i>odds and sods</i>
<i>Beat it!</i>	<i>(be) off your face</i>	<i>Old Glory</i>
<i>(be) the bee's knees</i>	<i>sweet Fanny Adams/FA</i>	<i>pull the other one / leg (it's got bells on)</i>
<i>(be) beside yourself</i>	<i>to a fare-thee-well</i>	<i>out (-) and (-) out</i>
<i>the Big Apple</i>	<i>buy the farm</i>	<i>peg it</i>
<i>sell sb a bill of goods (b/f)</i>	<i>chew the fat / rag</i>	<i>(be) (all) of a piece</i>
<i>give sb the bird (b/o)</i>	<i>(a bit of) how's your father</i>	<i>(be) a piece of cake / piss</i>
<i>(go) by the board(s)</i>	<i>put your foot in your mouth / in it (b/f)</i>	<i>take the piss</i>
<i>push the boat out</i>	<i>a French letter</i>	<i>come the raw prawn</i>
<i>Bob's your uncle</i>	<i>all gas and gaiters</i>	<i>part brass rags with</i>
<i>have a bone to pick with sb (b/f)</i>	<i>the ghost walks</i>	<i>(catch/have sb) bang / dead to rights</i>
<i>make no bones about it to boot</i>	<i>go bung (b/o)</i>	<i>take the rise out of sb</i>
<i>Box and Cox</i>	<i>the bitch goddess (b/f)</i>	<i>get your rocks off</i>
<i>shoot the breeze</i>	<i>cook sb's goose</i>	<i>row crosshanded</i>
<i>kick the bucket</i>	<i>out (-) of (-) hand</i>	<i>cut a / the rug</i>
<i>like nobody's business</i>	<i>knock into a cocked hat (b/f)</i>	<i>half seas over</i>
<i>butter wouldn't melt in sb's mouth (b/f)</i>	<i>(sb), eat your heart out!</i>	<i>serve sb right</i>
<i>by and by</i>	<i>red herring(s)</i>	<i>(be) at sixes and sevens</i>
<i>by and large</i>	<i>cock (-) a (-) hoop</i>	<i>so (-) and (-) so</i>
<i>not hold a candle to sb</i>	<i>(go) hell for leather</i>	<i>hop the stick / twig</i>
<i>(be) the cat's whiskers / pyjamas / miaow</i>	<i>cut no ice with sb (b/f)</i>	<i>such (-) and (-) such</i>
<i>bust sb's chops</i>	<i>a fine / pretty kettle of fish</i>	<i>to a T / tee</i>
<i>like the clappers (b/o)</i>	<i>lay out in lavender</i>	<i>take sb to task</i>
<i>clean sb's clock</i>	<i>swing the lead</i>	<i>tie one on</i>
<i>pop your clogs</i>	<i>pull sb's leg</i>	<i>touch (-) and (-) go</i>
<i>make sth out of whole cloth</i>	<i>trip the light fantastic</i>	<i>over the transom</i>
<i>heap coals of fire on sb's head</i>	<i>at loggerheads</i>	<i>cold turkey</i>
<i>come a cropper / purler</i>	<i>gird (up) your loins (b/f)</i>	<i>Uncle Sam</i>
<i>eat crow</i>	<i>So long</i>	<i>(be) off the wall</i>
<i>cut up rough</i>	<i>splice the mainbrace</i>	<i>and what have you</i>
<i>cut a dash</i>	<i>take the mickey</i>	<i>and what not</i>
<i>set sb / sth by the ears</i>	<i>not cut the mustard</i>	<i>wear the green willow</i>
	<i>(be) nip and tuck</i>	<i>-</i>

APPENDIX 8: ‘Most frequent idioms’ in both *Cambridge* & COBUILD

(*italics* indicates core idiom)

- | | |
|--|------------------------------------|
| 1. get your act together | 41. (be) over the moon |
| 2. get in on the act | 42. get on sb's nerves |
| 3. clear the air | 43. (speak) off the cuff |
| 4. jump on the bandwagon | 44. in a nutshell |
| 5. give someone the benefit of the doubt | 45. <i>out-and-out</i> |
| 6. (be) a bone of contention | 46. (be) plain sailing |
| 7. (be) out of the blue | 47. (be) in the pipeline |
| 8. (be) the bottom line | 48. take the plunge |
| 9. (be) in cold blood | 49. (be) off the record |
| 10. (be) in the dark | 50. set/put the record straight |
| 11. (be) down to earth | 51. red tape |
| 12. (be) easier said than done | 52. behind the scenes |
| 13. make ends meet | 53. from scratch |
| 14. make/pull a face | 54. (be) on the spot |
| 15. at first glance | 55. back to square one |
| 16. bear fruit | 56. on the spur of the moment |
| 17. give it a shot/whirl | 57. follow suit |
| 18. go against the grain | 58. in the short/medium/long term |
| 19. stick to your guns | 59. ram sth down sb's throat |
| 20. (be) a guinea pig | 60. keep sb on their toes |
| 21. get the hang of sth | 61. bite your tongue |
| 22. hold your own | 62. keep track |
| 23. let sb off the hook | 63. (have/get) the upper hand |
| 24. break the ice | 64. have your work cut out for you |
| 25. <i>(be) at loggerheads</i> | 65. (be) out of this world |
| 26. come to a head | 66. (be) up in the air |
| 27. draw the line at sth | 67. turn a blind eye to sth |
| 28. keep a low profile | 68. (be) over the top |
| 29. <i>take the mick/mickey</i> | 69. (be) out of your depth |
| 30. cross your mind | 70. (not) see eye to eye |
| 31. (be) on the cards | 71. not (be) the end of the world |
| 32. give/lend sb a hand | 72. get/come to grips with sth |
| 33. (be) run of the mill | 73. lose your head |
| 34. put your finger on sth | 74. break sb's heart |
| 35. give and take | 75. get your own back |
| 36. loud and clear | 76. fall into place |
| 37. bend/stretch the rules | 77. (be) on the right track |
| 38. play it safe | 78. bear/keep sth in mind |
| 39. (be) on top of sth | 79. stick/poke your nose in |
| 40. set/start the ball rolling | |

APPENDIX 9: Inter-rater reliability test 1

Pages **80, 160, 240, 320, 400** (*Cambridge International Dictionary of Idioms*)

1. cool it: 1 = compositional? Y? N?
2 = recognise untruth & reinterpret it? Y? N?
3 = some words literal/figurative? Y? N?
THEREFORE: CORE IDIOM - Y? N?
2. a cool customer: 1 = compositional? Y? N?
2 = recognise untruth & reinterpret it? Y? N?
3 = some words literal/figurative? Y? N?
THEREFORE: CORE IDIOM - Y? N?
3. a cool head: 1 = compositional? Y? N?
2 = recognise untruth & reinterpret it? Y? N?
3 = some words literal/figurative? Y? N?
THEREFORE: CORE IDIOM - Y? N?
4. be as cool as a cucumber: 1 = compositional? Y? N?
2 = recognise untruth & reinterpret it? Y? N?
3 = some words literal/figurative? Y? N?
THEREFORE: CORE IDIOM - Y? N?
5. keep your cool : 1 = compositional? Y? N?
2 = recognise untruth & reinterpret it? Y? N?
3 = some words literal/figurative? Y? N?
THEREFORE: CORE IDIOM - Y? N?
6. lose your cool: 1 = compositional? Y? N?
2 = recognise untruth & reinterpret it? Y? N?
3 = some words literal/figurative? Y? N?
THEREFORE: CORE IDIOM - Y? N?
7. play it cool: 1 = compositional? Y? N?
2 = recognise untruth & reinterpret it? Y? N?
3 = some words literal/figurative? Y? N?
THEREFORE: CORE IDIOM - Y? N?
8. fly the coop: 1 = compositional? Y? N?
2 = recognise untruth & reinterpret it? Y? N?
3 = some words literal/figurative? Y? N?
THEREFORE: CORE IDIOM - Y? N?
9. not be much cop: 1 = compositional? Y? N?
2 = recognise untruth & reinterpret it? Y? N?
3 = some words literal/figurative? Y? N?
THEREFORE: CORE IDIOM - Y? N?
10. copper-bottomed : 1 = compositional? Y? N?
2 = recognise untruth & reinterpret it? Y? N?
3 = some words literal/figurative? Y? N?
THEREFORE: CORE IDIOM - Y? N?
11. blot your copybook: 1 = compositional? Y? N?
2 = recognise untruth & reinterpret it? Y? N?
3 = some words literal/figurative? Y? N?
THEREFORE: CORE IDIOM - Y? N?

12. cut the (umbilical) cord: 1 = compositional? Y? N?
 2 = recognise untruth & reinterpret it? Y? N?
 3 = some words literal/figurative? Y? N?
 THEREFORE: CORE IDIOM - Y? N?
13. cordon bleu: 1 = compositional? Y? N?
 2 = recognise untruth & reinterpret it? Y? N?
 3 = some words literal/figurative? Y? N?
 THEREFORE: CORE IDIOM - Y? N?
14. to the core: 1 = compositional? Y? N?
 2 = recognise untruth & reinterpret it? Y? N?
 3 = some words literal/figurative? Y? N?
 THEREFORE: CORE IDIOM - Y? N?
15. around the corner: 1 = compositional? Y? N?
 2 = recognise untruth & reinterpret it? Y? N?
 3 = some words literal/figurative? Y? N?
 THEREFORE: CORE IDIOM - Y? N?
16. be backed into a corner: 1 = compositional? Y? N?
 2 = recognise untruth & reinterpret it? Y? N?
 3 = some words literal/figurative? Y? N?
 THEREFORE: CORE IDIOM - Y? N?

THEREFORE, ONLY _____ OF THE FIFTEEN WOULD BE JUDGED *CORE IDIOMS* BY MY CRITERIA.

Pages 80, 160, 240, 320, 400 (*Cambridge International Dictionary of Idioms*)

1. you can't keep a good woman down: 1 = compositional? Y? N?
 2 = recognise untruth & reinterpret it? Y? N?
 3 = some words literal/figurative? Y? N?
 THEREFORE: CORE IDIOM - Y? N?
2. kiss/say/wave goodbye to sth: 1 = compositional? Y? N?
 2 = recognise untruth & reinterpret it? Y? N?
 3 = some words literal/figurative? Y? N?
 THEREFORE: CORE IDIOM - Y? N?
3. a good-for-nothing: 1 = compositional? Y? N?
 2 = recognise untruth & reinterpret it? Y? N?
 3 = some words literal/figurative? Y? N?
 THEREFORE: CORE IDIOM - Y? N?
4. your goods and chattels: 1 = compositional? Y? N?
 2 = recognise untruth & reinterpret it? Y? N?
 3 = some words literal/figurative? Y? N?
 THEREFORE: CORE IDIOM - Y? N?
5. deliver/come up with the goods: 1 = compositional? Y? N?
 2 = recognise untruth & reinterpret it? Y? N?
 3 = some words literal/figurative? Y? N?
 THEREFORE: CORE IDIOM - Y? N?
6. a goody-goody: 1 = compositional? Y? N?
 2 = recognise untruth & reinterpret it? Y? N?
 3 = some words literal/figurative? Y? N?
 THEREFORE: CORE IDIOM - Y? N?

7. cook someone's goose: 1 = compositional? Y? N?
 2 = recognise untruth & reinterpret it? Y? N?
 3 = some words literal/figurative? Y? N?
 THEREFORE: CORE IDIOM - Y? N?
8. kills the goose that lays the golden egg: 1 = compositional? Y? N?
 2 = recognise untruth & reinterpret it? Y? N?
 3 = some words literal/figurative? Y? N?
 THEREFORE: CORE IDIOM - Y? N?
9. what's sauce/good for the goose (is sauce/good for the gander): 1 = compositional?
 Y? N?
 2 = recognise untruth & reinterpret it? Y? N?
 3 = some words literal/figurative? Y? N?
 THEREFORE: CORE IDIOM - Y? N?
10. a Gordian knot: 1 = compositional? Y? N?
 2 = recognise untruth & reinterpret it? Y? N?
 3 = some words literal/figurative? Y? N?
 THEREFORE: CORE IDIOM - Y? N?
11. Gordon Bennett: 1 = compositional? Y? N?
 2 = recognise untruth & reinterpret it? Y? N?
 3 = some words literal/figurative? Y? N?
 THEREFORE: CORE IDIOM - Y? N?

THEREFORE, ONLY _____ OF THE FIFTEEN WOULD BE JUDGED *CORE IDIOMS* BY MY CRITERIA.

Pages 80, 160, 240, 320, 400 (Cambridge International Dictionary of Idioms)

1. the luck of the draw: 1 = compositional? Y? N?
 2 = recognise untruth & reinterpret it? Y? N?
 3 = some words literal/figurative? Y? N?
 THEREFORE: CORE IDIOM - Y? N?
2. be down on your luck: 1 = compositional? Y? N?
 2 = recognise untruth & reinterpret it? Y? N?
 3 = some words literal/figurative? Y? N?
 THEREFORE: CORE IDIOM - Y? N?
3. be in/out of luck: 1 = compositional? Y? N?
 2 = recognise untruth & reinterpret it? Y? N?
 3 = some words literal/figurative? Y? N?
 THEREFORE: CORE IDIOM - Y? N?
4. have the luck of the devil: 1 = compositional? Y? N?
 2 = recognise untruth & reinterpret it? Y? N?
 3 = some words literal/figurative? Y? N?
 THEREFORE: CORE IDIOM - Y? N?
5. just my luck!: 1 = compositional? Y? N?
 2 = recognise untruth & reinterpret it? Y? N?
 3 = some words literal/figurative? Y? N?
 THEREFORE: CORE IDIOM - Y? N?
6. more by luck than judgement: 1 = compositional? Y? N?
 2 = recognise untruth & reinterpret it? Y? N?
 3 = some words literal/figurative? Y? N?
 THEREFORE: CORE IDIOM - Y? N?

7. no such luck!: 1 = compositional? Y? N?
 2 = recognise untruth & reinterpret it? Y? N?
 3 = some words literal/figurative? Y? N?
 THEREFORE: CORE IDIOM - Y? N?

8. push your luck: 1 = compositional? Y? N?
 2 = recognise untruth & reinterpret it? Y? N?
 3 = some words literal/figurative? Y? N?
 THEREFORE: CORE IDIOM - Y? N?

9. try your luck: 1 = compositional? Y? N?
 2 = recognise untruth & reinterpret it? Y? N?
 3 = some words literal/figurative? Y? N?
 THEREFORE: CORE IDIOM - Y? N?

10. your luck's in: 1 = compositional? Y? N?
 2 = recognise untruth & reinterpret it? Y? N?
 3 = some words literal/figurative? Y? N?
 THEREFORE: CORE IDIOM - Y? N?

11. get/strike it lucky: 1 = compositional? Y? N?
 2 = recognise untruth & reinterpret it? Y? N?
 3 = some words literal/figurative? Y? N?
 THEREFORE: CORE IDIOM - Y? N?

12. thank your lucky stars: 1 = compositional? Y? N?
 2 = recognise untruth & reinterpret it? Y? N?
 3 = some words literal/figurative? Y? N?
 THEREFORE: CORE IDIOM - Y? N?

13. you should be so lucky!: 1 = compositional? Y? N?
 2 = recognise untruth & reinterpret it? Y? N?
 3 = some words literal/figurative? Y? N?
 THEREFORE: CORE IDIOM - Y? N?

14. bring a lump to your throat: 1 = compositional? Y? N?
 2 = recognise untruth & reinterpret it? Y? N?
 3 = some words literal/figurative? Y? N?
 THEREFORE: CORE IDIOM - Y? N?

15. take your lumps: 1 = compositional? Y? N?
 2 = recognise untruth & reinterpret it? Y? N?
 3 = some words literal/figurative? Y? N?
 THEREFORE: CORE IDIOM - Y? N?

THEREFORE, ONLY _____ OF THE FIFTEEN WOULD BE JUDGED *CORE IDIOMS* BY MY CRITERIA.

Pages 80, 160, 240, 320, 400 (*Cambridge International Dictionary of Idioms*)

1. at the rate of knots: 1 = compositional? Y? N?
 2 = recognise untruth & reinterpret it? Y? N?
 3 = some words literal/figurative? Y? N?
 THEREFORE: CORE IDIOM - Y? N?

2. come the raw prawn: 1 = compositional? Y? N?
 2 = recognise untruth & reinterpret it? Y? N?
 3 = some words literal/figurative? Y? N?
 THEREFORE: CORE IDIOM - Y? N?

3. get a raw deal: 1 = compositional? Y? N?
2 = recognise untruth & reinterpret it? Y? N?
3 = some words literal/figurative? Y? N?
THEREFORE: CORE IDIOM - Y? N?
4. in the raw: 1 = compositional? Y? N?
2 = recognise untruth & reinterpret it? Y? N?
3 = some words literal/figurative? Y? N?
THEREFORE: CORE IDIOM - Y? N?
5. a ray of sunshine: 1 = compositional? Y? N?
2 = recognise untruth & reinterpret it? Y? N?
3 = some words literal/figurative? Y? N?
THEREFORE: CORE IDIOM - Y? N?
6. catch some/a few rays: 1 = compositional? Y? N?
2 = recognise untruth & reinterpret it? Y? N?
3 = some words literal/figurative? Y? N?
THEREFORE: CORE IDIOM - Y? N?
7. be/go out on the razzle: 1 = compositional? Y? N?
2 = recognise untruth & reinterpret it? Y? N?
3 = some words literal/figurative? Y? N?
THEREFORE: CORE IDIOM - Y? N?
8. razzle-dazzle: 1 = compositional? Y? N?
2 = recognise untruth & reinterpret it? Y? N?
3 = some words literal/figurative? Y? N?
THEREFORE: CORE IDIOM - Y? N?
9. reach for the moon/stars: 1 = compositional? Y? N?
2 = recognise untruth & reinterpret it? Y? N?
3 = some words literal/figurative? Y? N?
THEREFORE: CORE IDIOM - Y? N?
10. take it as read: 1 = compositional? Y? N?
2 = recognise untruth & reinterpret it? Y? N?
3 = some words literal/figurative? Y? N?
THEREFORE: CORE IDIOM - Y? N?
11. ready cash/money: 1 = compositional? Y? N?
2 = recognise untruth & reinterpret it? Y? N?
3 = some words literal/figurative? Y? N?
THEREFORE: CORE IDIOM - Y? N?
12. be ready to roll: 1 = compositional? Y? N?
2 = recognise untruth & reinterpret it? Y? N?
3 = some words literal/figurative? Y? N?
THEREFORE: CORE IDIOM - Y? N?
13. the real McCoy: 1 = compositional? Y? N?
2 = recognise untruth & reinterpret it? Y? N?
3 = some words literal/figurative? Y? N?
THEREFORE: CORE IDIOM - Y? N?
14. get real!: 1 = compositional? Y? N?
2 = recognise untruth & reinterpret it? Y? N?
3 = some words literal/figurative? Y? N?
THEREFORE: CORE IDIOM - Y? N?

15. you reap what you sow / As you sow, so shall you reap: 1 = compositional? Y? N?
 2 = recognise untruth & reinterpret it? Y? N?
 3 = some words literal/figurative? Y? N?
 THEREFORE: CORE IDIOM - Y? N?

THEREFORE, ONLY _____ OF THE FIFTEEN WOULD BE JUDGED *CORE IDIOMS* BY MY CRITERIA.

Pages 80, 160, 240, 320, 400 (Cambridge International Dictionary of Idioms)

1. get in touch (with): 1 = compositional? Y? N?
 2 = recognise untruth & reinterpret it? Y? N?
 3 = some words literal/figurative? Y? N?
 THEREFORE: CORE IDIOM - Y? N?
2. be out of touch: 1 = compositional? Y? N?
 2 = recognise untruth & reinterpret it? Y? N?
 3 = some words literal/figurative? Y? N?
 THEREFORE: CORE IDIOM - Y? N?
3. be in/out of touch with: 1 = compositional? Y? N?
 2 = recognise untruth & reinterpret it? Y? N?
 3 = some words literal/figurative? Y? N?
 THEREFORE: CORE IDIOM - Y? N?
4. lose your touch: 1 = compositional? Y? N?
 2 = recognise untruth & reinterpret it? Y? N?
 3 = some words literal/figurative? Y? N?
 THEREFORE: CORE IDIOM - Y? N?
5. be touch-and-go: 1 = compositional? Y? N?
 2 = recognise untruth & reinterpret it? Y? N?
 3 = some words literal/figurative? Y? N?
 THEREFORE: CORE IDIOM - Y? N?
6. touchy-feely: 1 = compositional? Y? N?
 2 = recognise untruth & reinterpret it? Y? N?
 3 = some words literal/figurative? Y? N?
 THEREFORE: CORE IDIOM - Y? N?
7. tough love: 1 = compositional? Y? N?
 2 = recognise untruth & reinterpret it? Y? N?
 3 = some words literal/figurative? Y? N?
 THEREFORE: CORE IDIOM - Y? N?
8. tough shit: 1 = compositional? Y? N?
 2 = recognise untruth & reinterpret it? Y? N?
 3 = some words literal/figurative? Y? N?
 THEREFORE: CORE IDIOM - Y? N?
9. a tough cookie: 1 = compositional? Y? N?
 2 = recognise untruth & reinterpret it? Y? N?
 3 = some words literal/figurative? Y? N?
 THEREFORE: CORE IDIOM - Y? N?
10. as tough as old boots/nails/shoe leather: 1 = compositional? Y? N?
 2 = recognise untruth & reinterpret it? Y? N?
 3 = some words literal/figurative? Y? N?
 THEREFORE: CORE IDIOM - Y? N?

11. hang tough: 1 = compositional? Y? N?
2 = recognize untruth & reinterpret it? Y? N?
3 = some words literal/figurative? Y? N?
THEREFORE: CORE IDIOM - Y? N?
12. a tour de force: 1 = compositional? Y? N?
2 = recognise untruth & reinterpret it? Y? N?
3 = some words literal/figurative? Y? N?
THEREFORE: CORE IDIOM - Y? N?
13. in tow: 1 = compositional? Y? N?
2 = recognise untruth & reinterpret it? Y? N?
3 = some words literal/figurative? Y? N?
THEREFORE: CORE IDIOM - Y? N?
14. be/go out on the town: 1 = compositional? Y? N?
2 = recognise untruth & reinterpret it? Y? N?
3 = some words literal/figurative? Y? N?
THEREFORE: CORE IDIOM - Y? N?
15. go to town on sth: 1 = compositional? Y? N?
2 = recognise untruth & reinterpret it? Y? N?
3 = some words literal/figurative? Y? N?
THEREFORE: CORE IDIOM - Y? N?

THEREFORE, ONLY _____ OF THE FIFTEEN WOULD BE JUDGED *CORE IDIOMS* BY MY CRITERIA.

APPENDIX 10: Inter-rater reliability test 2

Redefining ‘idioms’

Criteria have been previously established to define – or more accurately ‘describe’ – the present large collection of ‘idioms’, such as those found in idiom dictionaries. New criteria have been created to try to distinguish a smaller group of what will be called ‘core idioms’. ‘Idioms’ are therefore now divided into 3 categories – **figuratives**, **ONCEs**, and **core idioms**. These categories are explained below, along with examples of each category:

1. Figuratives – the meaning of the ‘idiom’ can be understood by recognising the ‘untruth’ in the MWU (multiword unit) and pragmatically reinterpreting it in a way that correctly explains the ‘idiom’.

Examples:	<i>a hidden agenda</i>	<i>be no picnic</i>
	<i>like a bat out of hell</i>	<i>put down roots</i>
	<i>the icing/frosting on the cake</i>	<i>be in the spotlight</i>
	<i>be in the dark</i>	<i>argue the toss</i>
	<i>grin and bear it</i>	<i>set the wheels in motion</i>
	<i>drop/fall into your lap</i>	<i>a zero-sum game</i>

2. ONCEs – only one word in the MWU is either not literal or non-compositional.

Examples:	<i>give someone the screaming abidabs</i>	<i>an Indian summer</i>
	<i>have a long face</i>	<i>rest on your laurels</i>
	<i>get a Guernsey</i>	<i>six of the best</i>

3. Core idioms – the MWU is both non-compositional (you can’t understand the meaning if you replace each lexical word with its own definition) and non-figurative (you can’t understand the meaning by recognising the untruth and pragmatically reinterpreting it, often with the aid of an image or figure).

Examples:	<i>have an axe to grind</i>	<i>red herring</i>
	<i>kick the bucket</i>	<i>out of hand</i>
	<i>cook somebody’s goose</i>	<i>take the mickey</i>

Of the three categories, it is ‘core idioms’ which will be the most difficult for learners to understand. The following two criteria will be used to define core idioms which, for the purposes of this study, must be MWUs (consisting of at least two words occurring elsewhere) that are:

- non-compositional
- non-figurative

In order to distinguish a ‘core idiom’ from other idioms as found in idiom dictionaries, the following test will be used for determining whether an MWU is a ‘core idiom’:

1. Is the meaning of the MWU retained if you replace each lexical word in the MWU with its own definition?

Y = compositional N = non-compositional

2. Is it possible to understand the meaning of the MWU by recognising the untruth and pragmatically reinterpreting it in a way that correctly explains the MWU?

Y = figurative N = non-figurative

3. Is there only one word in the MWU which is either not literal or non-compositional?

Y = non-relevant MWU N = relevant MWU

If we answer NO to these three questions, the MWU is a ‘core idiom’.

If we answer YES to any of these three questions, the MWU is not a ‘core idiom’ (may be literal, a figurative, or a ONCE; there will be some borderline cases).

However, in applying this test keep in mind that if we answer YES to the first question it will be ‘literal’ and thus the second and third questions are no longer applicable. Similarly, if we answer NO to the first question but YES to the second, the third question is no longer applicable as the construction will be ‘figurative’. Therefore, some ‘idioms’ could fall into either the ‘figuratives’ or the ‘ONCE’ categories, but if they are also figurative, it is not necessary to put them into the ONCE category. If, however, we answer NO to the first question, NO to the second but YES to the third, the construction will have only one word that is non-idiomatic and will become a ‘ONCE’.

Explaining the test

The steps will be outlined in a simple flow chart to show how the criteria can be applied to each MWU:

Criterion 1 =
compositional
(*red paint*) YES,
1 = compositional
(*red herring*) NO

Criterion 2 = figurative (*as good as gold*) YES,
2 = figurative
(*red herring*) NO

Criterion 3 = ONCE as only one part is ‘not literal’ or ‘non-compositional’
(*a long face*) YES
(*red herring*) NO
(*red herring*) is a ‘core idiom’

A number of other issues should be kept in mind:

- compound words and hyphenated words are considered single lexemes and will be excluded as they do not meet the criteria
- foreign phrase MWUs will be excluded as they do not meet the criteria (either the English translation would be a single lexeme, or a translation would explain the MWU)
- the results of the test do not necessarily match intuitions: even collocational fixed phrases such as *of course* would not normally be included in any ‘idiom’ category, and yet fit the criteria of being composed of at least two words which occur elsewhere and are ‘non-compositional, non-figurative’
- how easily someone can linguistically recognise and pragmatically reinterpret the untruth may vary according to their age, their knowledge of and exposure to these expressions, and so on. Some may react differently to an MWU such as *a dog in a manger*, one seeing it as figurative, the other as idiomatic. The difference may hinge on the word *manger*, the speaker who thought *a dog in a manger* was idiomatic associating the word purely with the Christmas story definition of *manger*, another seeing it primarily as a location for animals’ food.
- idiomatic expressions are seldom encountered outside of context (unless in lists) so contextual clues may be necessary to establish what a particular untruth might be referring to (e.g., *be like a dog with two tails* – the context should be compatible with ‘great happiness/excitement’, making the connection to the dog’s tail clear).

Testing the test

'Idioms' will now be taken from the same pages (pp. 150, 250, 350) of two different idiom dictionaries – the *Cambridge International Dictionary of Idioms* (1998) and the *Collins COBUILD Dictionary of Idioms* (1995).

<i>Cambridge</i>	Figurative?	ONCE?	C. I.	B/line?
(p.150)the full Monty				
allow/give sth full play				
at full stretch				
be as full as a boot / tick				
be full of yourself				
be full of beans				
be full of crap/ shit/ bull				
be full of holes				
be full of piss and vinegar				
be full of the joys of spring				
be in full cry				
be in full flow/ spate				
be in full swing				
not be the full quid				
come to a full stop				
come/go/turn full circle				
in full force				
(p.250) be meat and drink to sb				
be the meat in the sandwich				
deserve a medal				
a medallion man				
give sb a dose/ taste of their own medicine				
a meeting of minds				
a melting pot				
be a fully paid-up/a card-carrying member of sth				
commit sth to memory				
have a memory like an elephant				
jog your memory				
take a stroll/trip down memory lane				
the men in grey suits				
(p.350) a shrinking violet				
have a shufti				
shut your face/ gob/ mouth/ trap				
shuttle diplomacy				
fight shy of sth/ doing sth				
sick at heart				
be as sick as a dog				
be as sick as a parrot				
be sick and tired of/sick to death of sth/doing sth				
feel sick to your stomach				
be (on) the wrong side of 30/40, etc.				
be on the side of the angels				
be on the [expensive/ heavy/large etc.] side				
err on the side of caution				

know which side your bread is buttered (on)				
let the side down				

Collins COBUILD

Figurative? ONCE?

C.I.? B/line

(p.150) give sb the flick / the flick pass				
drop like flies				
There are no flies on sb				
a flight of fancy				
flip your lid/wig				
open the floodgates				
through the floor				
wipe the floor with sb				
(p.250) long on one thing& short on another				
a dirty / filthy / black look				
have never looked back				
throw/knock sb for a loop				
cut loose				
hang loose				
on the loose				
(p.350) on your side				
sunny side up				
to be on the safe side				
a brain like a sieve				
at first sight				
lose sight of sth				
out of sight, out of mind				
a sight for sore eyes				

Your job now is to apply the criteria and put the ‘idioms’ into the three new categories, or into a ‘borderline’ category if you’re not sure (if the criteria do not apply well enough to allow you to categorise the ‘idiom’).

APPENDIX 11: 'Idiom' articles and book chapters

surveyed for core idioms

1. Ackerman, B.P. (1982). *On comprehending idioms: Do children get the picture?*, *Journal of Experimental Child Psychology*, 33:439-454.
2. Adkins, P.G. (1968). Teaching idioms and figures of speech to non-native speakers of English. *Modern Language Journal*, 52:148-152.
3. Aijmer, K. (1996). *Conversational Routines in English, Convention and Creativity*. New York: Longman.
4. Alexander, R.J. (1987). Problems in understanding and teaching idiomaticity in English. *Anglistik und englischunterricht* 32:105-122.
5. Alexander, R.J. (1992). Fixed expressions, idioms and phraseology in recent learner's dictionaries. In *Euralex '92 Proceedings I-II*. Papers submitted to the 5th EURALEX International Congress on Lexicography (pp.35-41). Tampere, Finland: Publications of the Department of Translation Studies, University of Tampere.
6. Austin, P. and Kuiper, K. (1988). Constraints on coordinated idioms. *Te Reo, Journal of the Linguistic Society of New Zealand*, 31: 3-17.
7. Barkema, H. (1992). Idiomaticity in English NPs. In J. Aarts, P. de Haan, & N. Oostdijk. (Eds.). *English Language Corpora, Design, Analysis and Exploitation* (pp. 257-278). Rodopi: Amsterdam.
8. Bennett, T.J.A. (1988). *Aspects of English Colour Collocations and Idioms*. Heidelberg: Carl Winter.
9. Bobrow, S.A. and Bell, S.M. (1973). On catching onto idiomatic expressions. *Memory & Cognition* 1, 3: 343-346.
10. Bock, K.J. and Brewer, W.F. (1980). Comprehension and memory of the literal and figurative meaning of proverbs. *Journal of Psycholinguistic Research* 9, 1: 59-72.
11. Boers, F. and Demecheleer, M. (2001). Measuring the impact of cross-cultural differences on learners' comprehension of imageable idioms. *ELT Journal* 55, 3: 255-262.
12. Cacciari, C. and Tabossi, P. (1988). The comprehension of idioms. *Journal of Memory and Language*, 27: 668-683.
13. Cacciari, C. (1993). The place of idioms in a literal and metaphorical world. In Cacciari, C. and Tabossi, P. (Eds.). *IDIOMS: Processing, Structure, and Interpretation* (pp.27-55). Hillsdale, N.J.: Lawrence Erlbaum Associates.
14. Cacciari, C. and Glucksberg, S. (1991). Understanding idiomatic expressions: The contribution of word meanings. In Simpson, G.B. (Ed.). *Understanding Word and Sentence* (pp.217-240). North Holland: Elsevier Science Publishers B.V.
15. Cacciari, C. and Glucksberg, S. (1995). Imagining idiomatic expressions: Literal or figurative meaning? In M. Everaert, E-J.van der Linden, A. Schenk & R. Schreuder (Eds.). *Idioms: Structural and Psychological Perspectives* (pp.43-56).
16. Cermak, F. (1994). Idiomatics (pp.185-195). In P.A.Luelsdorff (Ed.). *The Prague School of Structural and Functional Linguistics, A Short Introduction*. Amsterdam / Philadelphia: John Benjamins.

17. Chafe, W.L. (1968). Idiomaticity as an anomaly of the Chomskyan paradigm. *Foundations of Language*, 4: 109-127.
18. Choul, J-C. (1980). Si mouve, ma non tropo: An inquiry into the non-metaphorical status of idioms and phrases. *Semiotics* (pp.89-98). Compiled by M. Herzfeld & M.D. Lenhart, Research Centre for Language and Semiotic Studies, Indiana University. New York: Plenum Press.
19. Cooper, T.C. (1998). Teaching idioms. *Foreign Language Annals* 31, 2:255-266.
20. Cooper, T.C. (1999). Processing of idioms by L2 learners of English, *TESOL Quarterly* 33, 2:233-262.
21. Coulmas, F. (1981). Idiomaticity as a problem of pragmatics. In H. Parret (Ed.). *Possibilities and Limitations of Pragmatics* (pp.139-151). Amsterdam: John Benjamins.
22. Cowie, A.P. (1992). Multiword lexical units and communicative language teaching. In P.J.L. Arnaud and H. Bejoint (Eds.). *Vocabulary and Applied Linguistics* (pp.1-12). Basingstoke: Macmillan.
23. Cowie, A.P. (1998). Phraseological dictionaries: Some east-west comparisons. In A.P. Cowie (Ed.). *Phraseology: Theory, Analysis, and Applications* (pp.209-228). Oxford: Clarendon Press.
24. Cutler, A. (1982). Idioms: The older, the colder. *Linguistic Inquiry*, 13:317-320.
25. Drew, P. and Holt, E. (1988). Complainable matters: The use of idiomatic expressions in making complaints. *Social Problems* 35, 4: 398-417.
26. Drew, P. and Holt, E. (1995). Idiomatic expressions and their role in the organization of topic transition in conversation. In M. Everaert, E-J.van der Linden, A. Schenk & R. Schreuder (Eds.). *IDIOMS: Structural and Psychological Perspectives* (pp. 117-132). Hillsdale, N.J.: Lawrence Erlbaum Associates.
27. Drew, P. and Holt, E. (1998). Figures of speech: Figurative expressions and the management of topic transition in conversation. *Language in Society*, 27: 495-522.
28. Estill, R.B. and Kemper, S. (1982). Interpreting idioms. *Journal of Psycholinguistic Research* 11, 6: 559-568.
29. Fernando, C. (1978). Towards a definition of idiom, its nature and function. *Studies in Language* 2, 3: 313-343.
30. Fernando, C. and Flavell, R. (1981). *On Idiom, Critical Views and Perspectives*. Exeter Linguistic Studies: University of Exeter.
31. Fernando, C. (1996). *Idioms and Idiomaticity*. Oxford: Oxford University Press.
32. Fillmore, C.J., Kay, P. & O'Connor, M.C. (1988). Regularity and idiomaticity in grammatical constructions: The case of let alone. *Language* 64, 3:501-538.
33. Forrester, M.A. (1995). Tropic implicature and context in the comprehension of idiomatic phrases. *Journal of Psycholinguistic Research* 24, 1:1-21.
34. Fraser, B. (1970). Idioms within a transformational grammar. *Foundations of Language*, 6, 1:22-42.
35. Gibbon, D. (1981). Idiomatic and functional variation: A case study of international amateur radio talk. *Language in Society* 10, 1:21-42.
36. Gibbs, R.W. Jr. (1985). On the process of understanding idioms. *Journal of Psycholinguistic Research* 14, 5:465-472.

37. Gibbs, R.W. Jr. (1986). Skating on thin ice: Literal meaning and understanding idioms in conversation. *Discourse Processes*, 9:17-30.
38. Gibbs, R.W. Jr. (1990). Psycholinguistic studies on the conceptual basis of idiomaticity. *Cognitive Linguistics* 1, 4:417-451.
39. Gibbs, R.W. Jr. (1992). What do idioms really mean? *Journal of Memory and Language*, 31: 485-506.
40. Gibbs, R.W. Jr. (1993). Why idioms are not dead metaphors. In C.Cacciari & P.Tabossi (Eds.). *IDIOMS, Processing, Structure and Interpretation* (pp.57-77). Hillsdale, N.J.: Lawrence Erlbaum Associates.
41. Gibbs, R.W. Jr. (1994). *The poetics of the mind*. Cambridge: Cambridge University Press.
42. Gibbs, R.W. Jr. (1995). Idiomaticity and human cognition. In M. Everaert, E-J.van der Linden, A. Schenk & R. Schreuder (Eds.). *IDIOMS: Structural and Psychological Perspectives* (pp. 97-116). Hillsdale, N.J.: Lawrence Erlbaum Associates.
43. Gibbs, R.W. Jr, & Gonzales, G.P. (1985). Syntactic frozenness in processing and remembering idioms. *Cognition*, 20:243-259.
44. Gibbs, R.W. Jr., & Nayak, N.P. (1989). Psycholinguistic studies on the syntactic behavior of idioms. *Cognitive Psychology*, 21:100-138.
45. Gibbs, R.W. Jr. & Nayak, N.P. (1991). Why idioms mean what they do. *Journal of Experimental Psychology*, 120, 1:93-95.
46. Gibbs, R.W. Jr, Nayak, N.P., Bolton, J.L. & Keppel, M.E. (1989). Speakers' assumptions about the lexical flexibility of idioms. *Memory & Cognition* 17, 1:58-68.
47. Gibbs, R.W. Jr, Nayak, N.P. & Cutting, C. (1989). How to kick the bucket and not decompose: Analyzability and idiom processing. *Journal of Memory and Language* 28:576-593.
48. Gibbs, R.W. Jr, and O'Brien, J.E. (1990). Idioms and mental imagery: The metaphorical motivation for idiomatic meaning. *Cognition*, 36:35-68.
49. Glaser, R. (1988). The grading of idiomaticity as a presupposition for a taxonomy of idioms. In W. Hullen & R. Schulze (Eds.). *Understanding the Lexicon* (pp.264-279). Tubingen: Max Niemeyer Verlag.
50. Glass, A.L. (1983). The comprehension of idioms. *Journal of Psycholinguistic Research* 12, 4:429-442.
51. Hatch, E. and Brown, C. (1995). *Vocabulary, Semantics and Language Education*. Cambridge: Cambridge University Press.
52. Holmes, V.L. and Moulton, M.R. (1993). Idioms in context: The Real McCoy. *TESL Reporter* 26, 2:41-45.
53. Irujo, S. (1986). A piece of cake: learning and teaching idioms. *ELT Journal* 40, 3:236-242.
54. Irujo, S. (1986). Don't put your leg in your mouth: Transfer in the acquisition of idioms in a second language. *TESOL Quarterly* 20, 2:287-304.
55. Irujo, S. (1993). Steering clear: Avoidance in the production of idioms. *IRAL* 31, 3:205-219.

56. Jackendoff, R. (1995). The boundaries of the lexicon. In M. Everaert, E-J.van der Linden, A. Schenk & R. Schreuder (Eds.). *IDIOMS: Structural and Psychological Perspectives* (pp.133-165). Hillsdale, N.J.: Lawrence Erlbaum Associates.
57. Jordan, R.R. (1995). The idiomatic English teacher (or keeping body and soul together). *Modern English Teacher* 4, 3:16.
58. Kellerman, E. (1979). Transfer and non-transfer: Where are we now. *Studies in Second Language Acquisition*, 2:37-57.
59. Kemper, S. (1986). Inferential processing and the comprehension of idioms. *Metaphor and Symbolic Activity* 1, 1:43-55.
60. Kerbel, D. & Grunwell, P. (1997). Idioms in the classroom: an investigation of language unit and mainstream teachers' use of idioms. *Child Language Teaching and Therapy* 3, 2:113-123.
61. Keysar, B. & Bly, B. (1995). Intuitions of the transparency of idioms: Can one keep a secret by spilling the beans? *Journal of Memory and Language* 34: 89-109.
62. Keysar, B. & Bly, B. (1999). Swimming against the current: Do idioms reflect conceptual structure? *Journal of Pragmatics*, 31: 1559-1578.
63. Kreuz, R.J. & Graesser, A.C. (1991). Aspects of idiom interpretation: Comment on Nayak and Gibbs. *Journal of Experimental Psychology* 120, 1:90-92.
64. Lattey, E. (1986). Pragmatic classification of idioms as an aid for the language learner. *IRAL* XXIV, 3: 217-233.
65. Lennon, P. (1998). Approaches to the teaching of idiomatic language. *IRAL*, XXXVI/I: 11-30.
66. Levorato, M.C. (1993). The acquisition of idioms and the development of figurative competence. In C.Cacciari and P.Tabossi (Eds.). *IDIOMS, Processing, Structure and Interpretation* (pp.101-128). Hillsdale, N.J.: Lawrence Erlbaum Associates.
67. Lorenz, E.K. (1977). Excuse me, but your idiom is showing. *The Reading Teacher*, October, 24-27.
68. Low, G.D. (1988). On teaching metaphor. *Applied Linguistics* 9, 2: 125-147.
69. Makkai, A. (1978). Idiomaticity as a language universal. In J.H. Greenberg, C.A. Ferguson, & E.A. Moravcsik (Eds.). *Universals of Human Language*, Volume 3: Word Structure (pp.401-448).. Stanford, California: Stanford University Press.
70. Makkai, A. (1975). The cognitive organization of idiomaticity: Rhyme or reason? *Georgetown University Working Papers of Language and Linguistics*, 11:10-29.
71. Malkiel, Y. (1959). Studies in irreversible binomials. *Lingua*, 8:113-160.
72. May, A.M. (1979). All the angles of idiom instruction. *The Reading Teacher*, March, 680-682.
73. Mayes, P. (1990). Quotation in spoken English. *Studies in Language* 14, 2:325-363.
74. McCarthy, M. (1990). *Vocabulary*. Oxford: Oxford University Press.
75. McCarthy, M. (1992). English idioms in use. *Revista Canaria de Estudios Ingleses*, 25: 55-65.
76. McCarthy, M. and Carter, R. (1994). *Language as Discourse: Perspectives for Language Teaching*. London and New York: Longman.

77. McGlone, M.S., Glucksberg, S., & Cacciari, C. (1994). Semantic productivity and idiom comprehension. *Discourse Processes*, 17:167-190.
78. Meier, H.H. (1975). On placing English idioms in lexis and grammar. *English Studies*, 231-244.
79. Mel'cuk, I. (1995). Phrasemes in language and phraseology in linguistics. In M. Everaert, E-J.van der Linden, A. Schenk & R. Schreuder (Eds.). *IDIOMS: Structural and Psychological Perspectives* (pp.167-232). Hillsdale, N.J.: Lawrence Erlbaum Associates.
80. Michiels, A. (1977). Idiomaticity in English. *Revue des Langues*, 43:184-199.
81. Moon, R. (1997). Vocabulary connections: multi-word items in English. In N. Schmitt and M. McCarthy (Eds.). *Vocabulary - Description, Acquisition and Pedagogy* (pp.40-63). Cambridge: Cambridge University Press.
82. Moon, R. (1992). There is reason for the roasting of eggs: a consideration of fixed expressions in native-speaker dictionaries. *Euralex '92 – Proceedings*, 493-502.
83. Moon, R. (1994). The analysis of fixed expressions in text. In M. Coulthard (Ed.). *Advances in Written Text Analysis* (pp.117-135). London and New York: Routledge.
84. Moon, R. (1998). Frequencies and forms of phrasal lexemes in English. In A.P. Cowie (Ed.). *Phraseology: Theory, Analysis, and Applications* (pp.79-100). Oxford: Clarendon Press.
85. Moulton, M. and Holmes, V. (1996). Teacher to teacher research in the ESL classroom. *TESL Reporter* 29, 1:34-40.
86. Mueller, R.A.G., and Gibbs, R.W. Jr. (1987). Processing idioms with multiple meanings. *Journal of Psycholinguistic Research* 16, 1:64-81.
87. Nattinger, J.R. (1980). A lexical phrase grammar for ESL. *TESOL Quarterly* XIV, 3:337-344.
88. Nayak, N.P. & Gibbs, R.W.Jr. (1990). Conceptual knowledge in the interpretation of idioms. *Journal of Experimental Psychology* 119, 3:315-330.
89. Needham, W.P. (1992). Limits on literal processing during idiom interpretation. *Journal of Psycholinguistic Research* 21, 1:1-16.
90. Nesi, H. (1995). A modern bestiary: a contrastive study of the figurative meanings of animal terms. *ELT Journal* 49, 3:272-278.
91. Newmeyer, F.J. (1974). The regularity of idiom behavior. *Lingua* 34:327-342.
92. Nicholas, T. (1995). Semantics of idiom modification. In M. Everaert, E-J.van der Linden, A. Schenk & R. Schreuder (Eds.). *IDIOMS: Structural and Psychological Perspectives* (pp.233-252). Hillsdale, N.J.: Lawrence Erlbaum Associates.
93. Nippold, M.A. (1998). *Later Language Development, The School-Age and Adolescent Years*. Second Edition. Austin, Texas: pro-ed.
94. Nippold, M.A. & Martin, S.T. (1989). Idiom interpretation in isolation versus context: A development study with adolescents. *Journal of Speech and Hearing Research* 32:59-66.
95. Nunberg, G, Sag, I.A. & Wasow, T. (1994). Idioms. *Language*, 70:491-534.
96. Odlin, T. (1991). Irish English idioms and language transfer. *English World-Wide* 12, 2:175-193.

97. Palmer, H.E. (1933). *Second Interim Report on English Collocations*. Tokyo: Kaitakusha.
98. Pawley, A. & Syder, F.H. (1983). Two puzzles for linguistic theory: Nativelike selection and nativelike fluency. In J.C. Richards & R. Schmidt (Eds.). *Language and Communication* (pp.191-216). New York: Longman.
99. Ponterotto, D. (1994). Metaphors we can learn by. *English Teaching Forum*, 33, 2:2-11.
100. Popiel, S.J. & McRae, K. (1988). The figurative and literal senses of idioms, or all idioms are not used equally. *Journal of Psycholinguistic Research* 17, 6:475-486.
101. Prinz, P.M. (1983). The development of idiomatic meaning in children. *Language and Speech* 26, 3:263-271.
102. Pulman, S.G. (1993). The recognition and interpretation of idioms. In C.Cacciari & P.Tabossi (Eds.). *IDIOMS, Processing, Structure and Interpretation* (pp.249-270). Hillsdale, N.J.: Lawrence Erlbaum Associates.
103. Reagan, R.T. (1987). The syntax of English idioms: Can the dog be put on?, *Journal of Psycholinguistic Research* 16, 5:417-441.
104. Roberts, M.H. (1944). The science of idiom: A method of inquiry into the cognitive design of language. *PMLA* LIX, 2:291-306.
105. Rose, J.H. (1978). Types of idiom. *Linguistics*, 203:55-62.
106. Ruhl, C. (1976). Idioms and data. *Lacus forum*, 456-466.
107. Schenk, A. (1995). The syntactic behavior of idioms. In M. Everaert, E-J.van der Linden, A. Schenk & R. Schreuder (Eds.). *IDIOMS: Structural and Psychological Perspectives* (pp.253-271). Hillsdale, N.J.: Lawrence Erlbaum Associates.
108. Schweigert, W.A. (1991). The muddy waters of idiom comprehension. *Journal of Psycholinguistic Research* 20, 4:305-314.
109. Schweigert, W.A. (1986). The comprehension of familiar and less familiar idioms. *Journal of Psycholinguistic Research* 15, 1:33-45.
110. Schweigert, W.A. & Moates, D.R. (1988). Familiar idiom comprehension. *Journal of Psycholinguistic Research* 17, 4:281-296.
111. Sornig, K. (1988). Idioms in language teaching. In W. Hullen & R. Schulze (Eds.). *Understanding the Lexicon, Meaning, Sense and World Knowledge in Lexical Semantics* (pp.280-290). Tübingen: Max Niemeyer.
112. Strassler, J. (1982). *Idioms in English. A Pragmatic Analysis*. Tübingen: Gunter Narr Verlag.
113. Swinney, D.A. (1979). The access and processing of idiomatic expressions. *Journal of Verbal Learning and Verbal Behavior*, 18:523-534.
114. Tabossi, P. & Zardon, F. (1993). The activation of idiomatic meaning in spoken language comprehension. In C.Cacciari & P.Tabossi (Eds.). *IDIOMS, Processing, Structure and Interpretation* (pp.145-162). Hillsdale, N.J.: Lawrence Erlbaum Associates.
115. Tabossi, P. and Zardon, F. (1995). The activation of idiomatic meaning. In M. Everaert, E-J.van der Linden, A. Schenk & R. Schreuder (Eds.). *Idioms: Structural and*

Psychological Perspectives (pp.273-282). Hillsdale, N.J.: Lawrence Erlbaum Associates.

116. Thompson, D. (1992). Television as a source of material for English dictionaries. *Euralex 92 – Proceedings*, 541-548.
117. Titone, D. & Connine, C. (1994). Descriptive norms for 171 idiomatic expressions: Familiarity, compositionality, predictability, and literality. *Metaphor and Symbolic Activity* 9, 4:247-270.
118. Titone, D. & Connine, C. (1999). On the compositional and noncompositional nature of idiomatic expressions. *Journal of Pragmatics*, 31:1655-1674.
119. Van Lancker, D. & Canter, G.J. (1981). Idiomatic vs literal interpretations of ditropically ambiguous sentences. *Journal of Speech and Hearing Research*, 46: 64-69.
120. Wallace, M.J. (1979). What is an idiom? An applied linguistics approach. In R.R.K.Hartmann (Ed.). *Dictionaries and their Users* (pp.63-70). University of Exeter: Exeter Linguistics Series.
121. Watson, J. (1998). Teaching idioms: Developing awareness of effect and appropriateness. *Modern English Teacher* 7, 1:15-19.
122. Weinreich, U. (1969). Problems in the analysis of idioms (pp.23-81). In J.Puhvel (Ed.). *Substance and Structure of Language*. Berkeley & Los Angeles: University of California Press.
123. Wood, M.M. (1986). *A Definition of Idiom*. Bloomington, Indiana: Indiana Linguistics Club.
124. Wray, A. and Perkins, M.R. (2000). The functions of formulaic language: An integrated model. *Language & Communication*, 20:1-28.
125. Yorio, C.A. (1980). Conventionalized language forms and the development of communicative competence. *TESOL Quarterly* XIV, 4:433-442.
126. Yorio, C.A. (1989). Idiomaticity as an indicator of second language proficiency. In K. Hytlenstam & L.K. Obler (Eds.). *Bilingualism across the Lifespan*. Aspects of acquisition, maturity, and loss (pp.55-72). Cambridge: Cambridge University Press.
127. Zeevat, H. (1995). Idiomatic blocking and the elsewhere principle. In M. Everaert, E-J.van der Linden, A. Schenk & R. Schreuder (Eds.). *Idioms: Structural and Psychological Perspectives*, pp.301-316.
128. Zgusta, L. (1967). Multiword lexical units. *Word*, 23:578-587.

BIBLIOGRAPHY

- Adkins, P. G. (1968). Teaching idioms and figures of speech to non-native speakers of English. *Modern Language Journal*, 52, 148-152.
- Aisenstadt, E. (1979). Collocability restrictions in dictionaries. *ITL: Review of Applied Linguistics*, 45(6), 71-74.
- Aisenstadt, E. (1981). Restricted collocations in English lexicology and lexicography. *ITL: Review of Applied Linguistics*, 53, 53-61.
- Alexander, R. (1978). Fixed expressions in English: a linguistic, psycholinguistic, sociolinguistic and didactic study (part 1). *Anglistik und Englischunterricht*, 6, 171-188.
- Alexander, R. (1984). Fixed expressions in English: reference books and the teacher. *ELTJ*, 38(2), 127-132.
- Alexander, R. J. (1987). Problems in understanding and teaching idiomaticity in English. *Anglistik und Englischunterricht*, 32, 105-122.
- Alexander, R. J. (1989). Fixed expressions, idioms and collocations revisited. In P. Meara (Ed.), *Beyond Words* (pp. 15-24): British Studies in Applied Linguistics 4.
- Ammer, C. (1992). *Have a Nice Day - No Problem! A Dictionary of Cliches*. New York: Plumb Books.
- Ammer, C. (1997). *The American Heritage Dictionary of Idioms*. Boston: Houghton Mifflin.
- Arnaud, P.J.L. & Savignon, S.J. (1997). Rare words, complex lexical units and the advanced learner. In J. Coady & T. Huckin (Eds.), *Second Language Vocabulary Acquisition* (pp.157-173). Cambridge: Cambridge University Press.
- Arnold, I. V. (1973). *The English Word*. Moscow: Vyssaja Skola.
- Auld, S. (2001). Exploring idioms usage. In A. Burns & H. de. S. Joyce (Eds.), *Teachers Voices: Teaching vocabulary* (Vol. 7) (pp. 120-128). Sydney: NCELTR.
- Bahns, J. (1993). Lexical collocations: a contrastive view. *ELT Journal*, 47(1), 56-61.
- Bahns, J., & Eldaw, M. (1993). Should we teach EFL students collocations? *System*, 21(1), 101-114.
- Bauer, L. (1998). When is a sequence of two nouns a compound in English? *English Language and Linguistics*, 2(1), 65-86.
- Bauer, L. & Huddleston, R. (2002). Lexical word-formation. In R. Huddleston & G.K. Pullum (Eds.), *The Cambridge Grammar of the English Language* (pp.1621-1721). Cambridge: Cambridge University Press.
- Becker, J. W. (1975). The phrasal lexicon. In B. Nash-Webber & R. Schank (Eds.), *Theoretical Issues in National Language Processing I*. Cambridge, Mass: Bolt, Beranek and Newman.
- Bell, J., & Gower, R. (1991). *Intermediate Matters*. Harlow, Essex: Longman.
- Bell, J., & Gower, R. (1992). *Upper Intermediate Matters*. Harlow, Essex: Longman.
- Bell, J., & Gower, R. (1995). *Pre-Intermediate Matters*. Harlow, Essex: Longman.
- Bell, J., & Gower, R. (1999). *Advanced Matters*. Harlow, Essex: Longman.

- Benson, M. (1985). Collocations and idioms. In R. Ilson (Ed.), *Dictionaries, Lexicography and Language Learning* (pp. 61-68): Pergamon Press/British Council.
- Boers, F. (2000a). Enhancing metaphoric awareness in specialised reading. *English for Specific Purposes*, 19(2), 137-147.
- Boers, F. (2000b). Metaphor awareness and vocabulary retention. *Applied Linguistics*, 21(4), 553-571.
- Boers, F. (2001). Remembering figurative idioms by hypothesising about their origin. *Prospect*, 16(3), 35-43.
- Boers, F., & Demecheleer, M. (1998). A cognitive semantic approach to teaching prepositions. *ELT Journal*, 52(3), 197-204.
- Boers, F., & Demecheleer, M. (2001). Measuring the impact of cross-cultural differences on learners' comprehension of imageable idioms. *ELT Journal*, 55(3), 255-262.
- Boers, F., & Eyckmans, J. (in press). Etymological elaboration as a strategy for learning figurative idioms. In P. Begaards & B. Laufer. (Eds.), *Vocabulary in a Second Language: Selection, Acquisition and Testing*. Amsterdam/Philadelphia: John Benjamins.
- Boers, F., & Littlemore, J. (2000). Cognitive style variables in participants' explanations of conceptual metaphors. *Metaphor and Symbol*, 15, 177-87.
- Bolinger, D. (1975). *Aspects of Language* (Second ed.). New York: Harcourt Brace Jovanovich.
- Bolinger, D. (1976). Meaning and memory. *Forum Linguisticum*, 1(1), 1-14.
- Burns, A., & de Silva Joyce, H. (2001). Investigating the teaching of vocabulary. In A. Burns, & H. de Silva Joyce. (Eds.), *Teachers Voices 7, Teaching Vocabulary*, (pp.vii-xviii). Sydney: NCELTR.
- Cacciari, C., & Glucksberg, S. (1991). Understanding idiomatic expressions: The contribution of word meanings. In G.B. Simpson (Ed.), *Understanding Word and Sentence* (pp.217-240). North Holland: Elsevier Science Publishers B.V.
- Cacciari, C., & Tabossi, P. (1988). The comprehension of idioms. *Journal of Memory and Language*, 27, 668-683.
- Cameron, L. (1999). Identifying and describing metaphor in spoken discourse data. In L. Cameron & G. Low (Eds.), *Researching and Applying Metaphor* (pp. 105-132). Cambridge: Cambridge University Press.
- Cameron, L., & Low, G. (1999). Metaphor. *Language Teaching*, 32, 77-96.
- Carter, R. (1987). *Vocabulary: Applied Linguistic Perspectives*. London: Routledge.
- Carter, R., & McCarthy, M. (1988). *Vocabulary and Language teaching*. New York: Longman.
- Carter, R., Goddard, A., Reah, D., Sanger, K., & Bowring, M. (1997). *Working with Texts*, A core book for language analysis. London: Routledge.
- Celce-Murcia, M., & Rosenweig, F. (1979). Teaching vocabulary in the ESL classroom. In M. Celce-Murcia & L. McIntosh. (Eds.), *Teaching English as a second or foreign language* (pp.241-257). Rowley, M.A.: Newbury House.
- Chafe, W. L. (1968). Idiomaticity as an anomaly in the Chomskyan paradigm. *Foundations of Language*, 4, 109-127.

- Charteris-Black, J. (2002). Second language figurative proficiency: a comparative study of Malay and English. *Applied Linguistics* 23: 104-33.
- Coady, J. (1997). L2 vocabulary acquisition through extensive reading. In J. Coady & T. Huckin (Eds.), *Second Language Vocabulary Acquisition* (pp.225-237). Cambridge: Cambridge University Press.
- Cooper, T. C. (1998). Teaching idioms. *Foreign Language Annals*, 31(2), 255-266.
- Cooper, T. C. (1999). Processing of idioms by L2 learners of English. *TESOL Quarterly*, 33(2), 233-262.
- Cornell, A. (1999). Idioms: An approach to identifying major pitfalls for learners. *IRAL*, 37(1), 1-22.
- Coulmas, F. (1981). Idiomaticity as a problem of pragmatics. In H. Parret, M. Sbisà & J. Verschueren (Eds.), *Possibilities and Limitations of Pragmatics* (pp. 139-151). Amsterdam: John Benjamins.
- Cowie, A. (1993). Getting to grips with phrasal verbs. *English Today* 36, 9(4), 38-41.
- Cowie, A. P. (1981). The treatment of collocations and idioms in learners' dictionaries. *Applied Linguistics*, 2(3), 223-235.
- Cowie, A. P. (1988). Stable and creative aspects of vocabulary use. In R. Carter & M. McCarthy (Eds.), *Vocabulary and Language Teaching* (pp. 126-137). London: Longman.
- Cowie, A. P. (1992). Multiword lexical units and communicative language teaching. In P. J. L. Arnaud & H. Bejoint (Eds.), *Vocabulary and Applied Linguistics* (pp. 1-12). Basingstoke: MacMillan.
- Cowie, A. P. (1998). Phraseological dictionaries: Some East-West comparisons. In A.P. Cowie (Ed.), *Phraseology, Theory, Analysis, and Applications* (pp.209-228). Oxford: Clarendon Press.
- Cowie, A. P., & Mackin, R. (1975). *Oxford Dictionary of Current Idiomatic English, Volume 1: Verbs with Prepositions and Particles*. London: Longman.
- Cowie, A. P., & Mackin, R. (1993). *Oxford Dictionary of Phrasal Verbs*. Oxford: Oxford University Press.
- Cowie, A. P., Mackin, R., & McCaig, I. R. (1983). *Oxford Dictionary of Current Idiomatic English: Phrase, Clause and Sentence Idioms* (Vol. 2). Oxford: Oxford University Press.
- Cowie, A. P., Mackin, R., & McCaig, I. R. (1993). *Oxford Dictionary of English Idioms*. Oxford: Oxford University Press.
- Cruse, D. (1986). *Lexical Semantics*. Cambridge: Cambridge University Press.
- Cryer, M. (2001). *Curious Questions*. Auckland: Harper Collins.
- Cunningham, S., & Moor, P. (1998). *Cutting Edge Intermediate*. Harlow, Essex: Longman.
- Cunningham, S., & Moor, P. (1999). *Cutting Edge Upper Intermediate*. Harlow, Essex: Longman.
- Deignan, A. D. (1995). *Collins Cobuild Guides 7: Metaphor*. London: Harper Collins.

- Deignan, A., Gabrys, D., & Solska, A. (1997). Teaching English metaphors using cross-linguistic awareness-raising activities. *ELT Journal*, 51(4), 352-360.
- Doff, A., & Jones, C. (1994). *Language in use - Intermediate*. Cambridge: Cambridge University Press.
- Doff, A., & Jones, C. (1997). *Language in use - Upper-Intermediate*. Cambridge: Cambridge University Press.
- Doff, A., & Jones, C. (1999). *Language in use - Beginner*. Cambridge: Cambridge University Press.
- Doff, A., & Jones, C. (2000). *Language in use - Pre-Intermediate*. Cambridge: Cambridge University Press.
- Donald, G. (1994). *The Dictionary of Modern Phrase*. London: Simon & Schuster.
- Drew, P., & Holt, E. (1988). Complaining matters: the use of idiomatic expressions in making complaints. *Social Problems*, 35(4), 398-417.
- Drew, P., & Holt, E. (1995). Idiomatic expressions and their role in the organization of topic transition in conversation. In M. Everaert & et al (Eds.), *IDIOMS: Structural and Psychological Perspectives* (pp. 117-131). Hillsdale, N. J.: Lawrence Erlbaum Associates.
- Drew, P., & Holt, E. (1998). Figures of speech: Figurative expressions and the management of topic transition in conversation. *Language in Society*, 27(4), 495-522.
- Ellis, N. C. (2001). Memory for language. In P. Robinson (Ed.), *Cognition and Second Language Instruction* (pp.33-68). Cambridge: Cambridge University Press.
- Ellis, R. (1994). *The Study of Second Language Acquisition*. Oxford: Oxford University Press.
- Ellis, R. (1997). *SLA Research and Language Teaching*. Oxford: Oxford University Press.
- Farghal, M., & Obiedat, H. (1995). Collocations: a neglected variable in EFL. *IRAL*, XXXIII(4), 315-331.
- Fernando, C. (1978). Towards a definition of idiom, its nature and function. *Studies in Language*, 2(3), 313-343.
- Fernando, C. (1996). *Idioms and Idiomaticity*. Oxford: Oxford University Press.
- Fernando, C., & Flavell, R. (1981). *On Idiom: Critical Views and Perspectives*. Exeter: University of Exeter.
- Fillmore, C., Kay, P., & O'Connor, M. (1988). Regularity and idiomaticity in grammatical constructions: The case of *let alone*. *Language*, 64, 501-538.
- Firth, J. R. (1957). Modes of Meaning, *Papers in Linguistics*. Oxford: Oxford University Press.
- Flavell, L. & Flavell, R. (1992). *Dictionary of Idioms and their Origins*. London: Kyle Cathie.
- Flavell, L. & Flavell, R. (1993). *Dictionary of Proverbs and their Origins*. London: Kyle Cathie.
- Fraser, B. (1970). Idioms within a transformational grammar. *Foundations of Language*, 6(1), 22-42.
- Gazdar, G., Klein, E., Pullam, G., & Sag, I. (1985). *Generalized Phrase Structure Grammar*. Oxford: Basil Blackwell.

- Gibbs, R. W. (1990). Psycholinguistic studies on the conceptual basis of idiomaticity. *Cognitive Linguistics*, 1(4), 417-451.
- Gibbs, R.W. (1992). What do idioms really mean? *Journal of Memory and Language*, 31, 485-506.
- Gibbs, R. W. (1994). *The poetics of the mind - Figurative language, thought and understanding*. Cambridge: Cambridge University Press.
- Gibbs, R. W. (1997). Taking metaphor out of our heads and putting it into the cultural world. In R.W. Gibbs & R. Schulze (Eds.), *Understanding the Lexicon* (pp. 145-166). Tübingen: Max Niemeyer Verlag.
- Gibbs, R. W., & Nayak, N. (1989). Psycholinguistic studies on the syntactic behavior of idioms. *Cognitive Psychology*, 21, 100-138.
- Gibbs, R. W., Nayak, N., Bolton, J., & Keppel, M. (1989). Speakers' assumptions about the lexical flexibility of idioms. *Memory and Cognition*, 17, 58-68.
- Gibbs, R. W., Nayak, N., & Cutting, C. (1989). How to kick the bucket and not decompose: Analyzability and idiom processing. *Journal of Memory and Language*, 28, 576-593.
- Glaser, R. (1988). The grading of idiomaticity as a presupposition for a taxonomy of idioms. In W. Hüllen & R. Schulze (Eds.), *Understanding the Lexicon* (pp.264-279). Tübingen: Max Niemeyer Verlag.
- Goatly, A. (1997). *The Language of Metaphors*. London: Routledge.
- Goodale, M. (1995). *Collins COBUILD IDIOMS Workbook*. London: Harper Collins Publishers.
- Grabe, W. & Stoller, F. (1997). Reading and vocabulary development in a second language: A case study. In J. Coady & T. Huckin (Eds.), *Second Language Vocabulary Acquisition* (pp.98-122). Cambridge: Cambridge University Press.
- Greenall, S. (1994). *Reward - Pre-Intermediate*. Oxford: Macmillan Heinemann ELT.
- Greenall, S. (1995). *Reward - Intermediate*. Oxford: Macmillan Heinemann ELT.
- Greenall, S. (1996). *Reward - Upper-Intermediate*. Oxford: Macmillan Heinemann ELT.
- Gulland, D.M., & Hinds-Howell, D. (1986). *The Penguin Dictionary of English Idioms*. Hammondsworth: Penguin Books.
- Hatch, E., & Brown, C. (1995). *Vocabulary, Semantics, and Language Education*. Cambridge: Cambridge University Press.
- Hockett, C. F. (1958). *A Course in Modern Linguistics*. New York: Macmillan.
- Holmes, V. L., & Moulton, M. R. (1993). Idioms in context: The real McCoy. *TESL Reporter*, 26(2), 41-45.
- Honeyfield, J. (1977). Word frequency and the importance of context in vocabulary learning. *RELC Journal*, 8(2), 35-42.
- Howarth, P. (1996). *Phraseology in English Academic Writing. Some implications for language learning and dictionary making*. Tübingen: Max Niemeyer Verlag.
- Howarth, P. (1998). Phraseology and second language proficiency. *Applied Linguistics*, 19(1), 24-44.

- Hulstijn, J.H. (1997). Mnemonic methods in foreign language vocabulary learning. In J. Coady & T. Huckin (Eds.), *Second Language Vocabulary Acquisition* (pp.203-224). Cambridge: Cambridge University Press.
- Irujo, S. (1986a). Don't put your leg in your mouth: transfer in the acquisition of idioms in a second language. *TESOL Quarterly*, 20(2), 287-304.
- Irujo, S. (1986b). A piece of cake: learning and teaching idioms. *ELT Journal*, 40(3), 236-242.
- Irujo, S. (1993). Steering clear: Avoidance in the production of idioms. *IRAL*, 31(3), 205-219.
- Janssen, T. M. V. (Ed.). (1994). *Compositionality of meaning* (Vol. 2). Oxford: Pergamon Press.
- Johnson, J. M. (1996). Metaphor interpretation by second language learners: Children and adults. *The Canadian Modern Language Review*, 53(1), 219-241.
- Katz, J., & Postal, P. (1963). Semantic interpretation of idioms and sentences containing them. *MIT Research Laboratory of Electronics, Quarterly Progress Report*, 70, 275-282.
- Kay, P. (1993). *Words and Grammar of Context*. Stanford USA: CSLI Publications.
- Kennedy, G. D. (1987). Expressing temporal frequency in academic English. *TESOL Quarterly*, 21(1), 69-85.
- Kerbel, D., & Grunwell, P. (1997). Idioms in the classroom: An investigation of language unit and mainstream teachers' use of idioms. *Child Language Teaching and Therapy*, 13(2), 113-123.
- Kirkpatrick, E. M., & Schwarz, C. M. (Eds.). (1993). *The Wordsworth Dictionary of Idioms*. Ware, Hertfordshire: Wordsworth Editions Ltd.
- Kjellmer, G. (1994). *A Dictionary of English Collocations, Based on the Brown Corpus*. Oxford: Pergamon Press.
- Koester, A. J. (2000). The role of idioms in negotiating workplace encounters. In J. Trappes-Lomax (Ed.), *Change and Continuity in Applied Linguistics* (pp. 169-183). Clevedon: Multilingual Matters.
- Kovecses, G. (1997). Metaphor: Does it constitute or reflect cultural models? In R.W. Gibbs & G. J. Steen (Eds.), *Metaphor in Cognitive Linguistics* (pp. 167-188). Amsterdam: John Benjamins.
- Kovecses, Z., & Szabo, P. (1996). Idioms: a view from cognitive semantics. *Applied Linguistics*, 17(3), 326-355.
- Lakoff, G. (1986) The meaning of literal. *Metaphor and Symbolic Activity*, 1(4), 291-296.
- Lakoff, G. (1987). *Women, fire, and dangerous things*. Chicago: Chicago University Press.
- Lakoff, G., & Johnson, M. (1980). *Metaphors We Live By*. Chicago: University of Chicago Press.
- Lakoff, G., & Turner, M. (1989). *More than Cool Reason: A Field Guide to Poetic Metaphor*. Chicago: University of Chicago Press.
- Langacker, R. (1986). *Foundations of Cognitive Grammar* (Vol. 1). Stanford: Stanford University Press.
- Laufer, B. (1990). Why are some words more difficult than others? - Some intralexical factors that affect the learning of words. *IRAL*, 28(4), 293-307.

- Laufer, B. (1992). *Corpus-based versus lexicographer examples in comprehension and production of new words*. Paper presented at the Euralex 92 Proceedings I-II, Papers submitted to the 5th EURALEX International Congress on Lexicography, Tampere, Finland.
- Lazar, G. (1996). Using figurative language to expand students' vocabulary. *ELT Journal*, 50(1), 43-51.
- Lazar, G. (2003). *Meanings and Metaphors*. Cambridge: Cambridge University Press.
- Leech, G., Rayson, P., & Wilson, A. (2001). *Word Frequencies in Written and Spoken English, based on the British National Corpus*. Harlow and London: Longman.
- Lennon, P. (1998). Approaches to the teaching of idiomatic language. *IRAL*, XXXVI(1), 11-30.
- Lewis, M. (1997). Pedagogical implications of the lexical approach. In J. Coady & T. Huckin (Eds.), *Second Language Vocabulary Acquisition* (pp.255-270). Cambridge: Cambridge University Press.
- Lewis, M. (1993). *The Lexical Approach*. Hove, UK: Language Teaching Publications.
- Lewis, M. (Ed.) (2000). *Teaching Collocation, Further developments in the Lexical Approach*. Hove, UK: Language Teaching Publications.
- Lipka, L. (1991). *An Outline of English Lexicology*. Tübingen: Max Niemeyer.
- Low, G. D. (1988). On teaching metaphor. *Applied Linguistics*, 9(2), 125-147.
- Low, G. (1999). Validating metaphor research projects. In L. Cameron & G. Low (Eds.), *Researching and Applying Metaphor* (pp. 48-65). Cambridge: Cambridge University Press.
- Lyons, J. (1977). *Semantics I* (Vol. 1). Cambridge: Cambridge University Press.
- Lyons, J. (1995). *Linguistic Semantics, An Introduction*. Cambridge: Cambridge University Press.
- MacLennan, C. (1993). Metaphor in the language classroom: A case for change. *ILEJ*, 10, 137-152.
- MacLennan, C. H. G. (1994). Metaphors and prototypes in the learning and teaching of grammar and vocabulary. *IRAL*, XXXII(2), 98-110.
- Makkai, A. (1965). *Idiom Structure in English*. Unpublished Dissertation, Yale.
- Makkai, A. (1969). The two idiomaticity areas in English and their membership: A stratificational view. *Linguistics*, 50, 44-58.
- Makkai, A. (1972). *Idiom Structure in English*. The Hague: Mouton.
- Makkai, A. (1975). The cognitive organization of idiomaticity: rhyme or reason? *Georgetown University Working Papers of Language and Linguistics*, 11, 10-29.
- Malkiel, Y. (1959). Studies in irreversible binomials. *Lingua*, 8, 113-160.
- Manser, M. H. (1997). *A Dictionary of Everyday Idioms* (Second ed.). London: Prentice Hall Europe ELT.
- McCarthy, M. J. (1984). A new look at vocabulary in EFL. *Applied Linguistics*, 5(1), 12-22.
- McCarthy, M. J. (1991). *Discourse Analysis for Language Teachers*. Cambridge: Cambridge University Press.

- McCarthy, M. J. (1992). English idioms in use. *Revista Canaria de Estudios Ingleses*, 25, 55-65.
- McCarthy, M. J. (1998). *Spoken language and applied linguistics*. Cambridge: Cambridge University Press.
- McCarthy, M., & Carter, R. (1994). *Language as Discourse: Perspectives for Language Teaching*. London: Longman.
- McCarthy, M., & O'Dell, F. (1994). *English Vocabulary in Use, Upper-Intermediate & Advanced*. Cambridge: Cambridge University Press.
- McCarthy, M., & O'Dell, F. (1999). *English Vocabulary in Use, Elementary*. Cambridge: Cambridge University Press.
- McCarthy, M., & O'Dell, F. (2002). *English Idioms in Use*. Cambridge: Cambridge University Press.
- McCarthy, M. J. & Walter, E. (1997). *Cambridge International Dictionary of Phrasal Verbs*. Cambridge: Cambridge University Press.
- McCarthy, M. J., & Walter, E. (1998). *Cambridge International Dictionary of Idioms*. Cambridge: Cambridge University Press.
- Meara, P. (1980). Vocabulary acquisition: A neglected aspect of language learning. *Language Teaching and Linguistics Abstracts*, 13, 221-246.
- Mel'cuk, I. A. (1960). O terminach 'ustojcivost' i 'idiomaticnost'. *Voprosy jazykoznanija*, 4, 73-80.
- Mel'cuk, I. A. (1988). Semantic description of lexical units in an explanatory combinatorial dictionary: Basic principles and heuristic criteria. *International Journal of Lexicography*, 1(3), 165-188.
- Mel'cuk, I. A., Arbatchewsky-Jumarie, N., Elnitsky, L., Iordanskaja, L., & Lessard, A. (1984). *Dictionnaire explicatif et combinatoire du français contemporain: Recherches lexicosemantiques I*. Montreal: Les Presses de l'Universite de Montreal.
- Michiels, A. (1977). Idiomaticity in English. *Revue des Langues*, 43, 184-197.
- Mitchell, T. F. (1971). Linguistic 'goings-on': collocations and other lexical matters arising on the syntagmatic record. *Archivum Linguisticum*, 2, 35-69.
- Moon, R. (1997). Vocabulary connections: multi-word items in English. In N. Schmitt & M. McCarthy (Eds.), *Vocabulary, Description, Acquisition and Pedagogy* (pp.40-63). Cambridge: Cambridge University Press.
- Moon, R. (1998a). *Fixed Expressions and Idioms in English: A Corpus-Based Approach*. Oxford: Clarendon Press.
- Moon, R. (1998b). Frequencies and forms of phrasal lexemes in English. In A. P. Cowie (Ed.), *Phraseology, Theory, Analysis, and Applications* (pp. 79-100). Oxford: Clarendon Press.
- Nation, I. S. P. (1990). *Teaching and Learning Vocabulary*. Massachusetts: Newbury House.
- Nation, P. (2000). Learning vocabulary in lexical sets: dangers and guidelines. *TESOL Journal*, Summer. 6-10.
- Nation, I. S. P. (2001). *Learning Vocabulary in Another Language*. Cambridge: Cambridge University Press.

- Nation, I.S.P. & Coady, J. (1988). Vocabulary and reading. In R. Carter & M. McCarthy (Eds.), *Vocabulary and language teaching* (pp.97-110). London: Longman.
- Nation, P. & Newton, J. (1997). Teaching vocabulary. In J. Coady & T. Huckin (Eds.), *Second Language Vocabulary Acquisition* (pp.238-254). Cambridge: Cambridge University Press.
- Nattinger, J. R., & DeCarrico, J. S. (1992). *Lexical Phrases and Language Teaching*. Oxford: Oxford University Press.
- Naylor, L. (2002a). *Join the Club 1*, Idioms for Academic and Social Success. USA: McGraw-Hill Contemporary.
- Naylor, L. (2002b). *Join the Club 2*, Idioms for Academic and Social Success. USA: McGraw-Hill Contemporary.
- Newmeyer, F. (1972). The insertion of idioms, *Papers from the English Regional Meeting, Chicago Linguistic Society* (pp. 294-302). Chicago: Chicago Linguistic Society.
- Newmeyer, F. J. (1974). The regularity of idiom behavior. *Lingua*, 34, 327-342.
- Nunberg, G. (1978). *The pragmatics of reference*. Bloomington: Indiana University Linguistics Club.
- Nunberg, G., Sag, I., & Wasow, T. (1994). Idioms. *Language*, 70, 491-534.
- Ortony, A. (1984). Understanding figurative language. In Pearson (Ed.), *Handbook of Reading Research* (pp. 453-470). New York: Longman.
- Oxford Idioms Dictionary for learners of English*. (2001). Oxford: Oxford University Press.
- Palmer, H. E. (1933). *Second Interim Report on English Collocations*. Tokyo: Kaitakusha.
- Partee, B. H., ter Meulen, A., & Wall, R. E. (1990). *Mathematical Methods in Linguistics*. Dordrecht, The Netherlands: Kluwer Academic Publishers.
- Partridge, E. (1978). *You Have a Point There*. A guide to punctuation and its allies. London: Routledge & Kegan Paul Ltd.
- Pawley, A., & Syder, F. H. (1983). Two puzzles for linguistic theory: nativelike selection and nativelike fluency. In J. C. Richards & R. W. Schmidt (Eds.), *Language and Communication* (pp. 191-225). New York: Longman.
- Pollio, H. R., Barlow, J. M., Fine, H. J., & Pollio, M. R. (1977). *Psychology and the poetics of growth: Figurative language in psychology, psychotherapy and education*. Hillsdale, N J: Lawrence Erlbaum Associates.
- Popiel, S. J., & McRae, K. (1988). The figurative and literal senses of idioms, or all idioms are not used equally. *Journal of Psycholinguistic Research*, 17(6), 475-487.
- Reagan, R. T. (1987). The syntax of English idioms: can the dog be put on? *Journal of Psycholinguistic Research*, 16(3), 417-441.
- Redman, S. (1997). *English Vocabulary in Use, Pre-Intermediate & Intermediate*. Cambridge: Cambridge University Press.
- Roberts, M. H. (1944). The science of idiom: A method of inquiry into the cognitive design of language. *PMLA*, LIX, 291-306.
- Ruhl, C. (1975). Kick the bucket is not an idiom. *Interfaces*, 4, 2-4.

- Schmidt, R. (1990). The role of consciousness in second language learning. *Applied Linguistics*, 11:129-158.
- Schmidt, R., & Frota, S. (1986). Developing basic conversational ability in a second language : A case-study of an adult learner. In R. Day (Ed.), *Talking to Learn: Conversation in Second Language Acquisition* (pp.237-326). Rowley, Mass.: Newbury House.
- Schweigert, W. A., & Moates, D. R. (1988). Familiar idiom comprehension. *Journal of Psycholinguistic Research*, 17(4), 281-296.
- Seidl, J., & McMordie, W. (1988). *English Idioms* (Fifth ed.). Oxford: Oxford University Press.
- Seymour, G. (2001). *The Untouchable*. London: Bantam.
- Shaptozhvili, S. (2002). Vocabulary practice games. *English Teaching Forum*, 40(2), 34-37.
- Simpson, R. & Mendis, D. (2003). A corpus-based study of idioms in academic speech. *TESOL Quarterly*, 37, 3:419-441.
- Sinclair, J. (1987). Collocation: a progress report. In R. Steele & T. Threadgold (Eds.), *Language Topics: Essays in honour of Michael Halliday* (Vol. 2, pp. 319-331). Amsterdam: John Benjamins.
- Sinclair, J. (1991a). *Corpus, Concordance, Collocation*. Oxford: Oxford University Press.
- Sinclair, J. (1991b). *Collins English Dictionary*. Glasgow: HarperCollins Publishers.
- Sinclair, J. (1994). *Collins English Dictionary*. Glasgow: HarperCollins Publishers.
- Sinclair, J., & Moon, R. (1989). *Collins COBUILD Dictionary of Phrasal Verbs*. London: Harper Collins Publishers.
- Sinclair, J., & Moon, R. (1995). *Collins COBUILD Dictionary of Idioms*. London: Harper Collins Publishers.
- Soars, L., & Soars, J. (1998a). *New Headway Pre-Intermediate*. Oxford: Oxford University Press.
- Soars, L., & Soars, J. (1998b). *New Headway Upper-Intermediate*. Oxford: Oxford University Press.
- Soars, L., & Soars, J. (2000a). *New Headway Elementary*. Oxford: Oxford University Press.
- Soars, L., & Soars, J. (2000b). *New Headway Intermediate*. Oxford: Oxford University Press.
- Sokman, A. J. (1992). Students as vocabulary generators. *TESOL Journal*, 1, 16-18.
- Speake, J. (1999). *The Oxford Dictionary of Idioms*. New York: Oxford University Press.
- Stern, K. (1998). *Longman Idioms Dictionary*. Harlow, Essex: Longman.
- Strassler, J. (1982). *Idioms in English: A Pragmatic Analysis*. Tübingen: Gunter Narr Verlag.
- Thomas, B. J. (1989). *Advanced Vocabulary & Idiom*. London: Longman.
- Thomas, B. J. (1995). *Advanced Vocabulary & Idiom* (New ed.). London: Longman.
- Titone, D. A., & Connine, C. M. (1999). On the compositional and noncompositional nature of idiomatic expressions. *Journal of Pragmatics*, 31, 1655-1672.
- Trask, R. L. (1997). *The Penguin Guide to Punctuation*. London: Penguin.

- Waretini, J. (2002). Opening a can of worms: The prevalence of idiomatic and colloquial language encountered by the ESOL student in New Zealand. *TESOLANZ Newsletter*, 11(1), 12.
- Waring, R. (1997). A study of receptive and productive learning from word cards. *Studies in Foreign Languages and Literature*, 21(1), 94-114.
- Watson, J. (1998). Teaching idioms: Developing awareness of effect and appropriateness. *Modern English Teacher*, 7(1), 15-19.
- Weinreich, U. (1969). Problems in the analysis of idioms. In J. Puhvel (Ed.), *Substance and Structure of Language* (pp. 23-81). Berkeley & Los Angeles: University of California Press.
- Why Do We Say It?* (1985). The stories behind the words, expressions and clichés we use. Edison, N.J.: Castle Books.
- Wood, M. M. (1981). *A Definition of Idiom*. Unpublished Master of Arts, University of Birmingham, Birmingham.
- Wray, A. (2000). Formulaic sequences in second language teaching: Principle and practice. *Applied Linguistics*, 21(4), 463-489.
- Wray, A., & Perkins, M. R. (2000). The function of formulaic language: an integrated model. *Language and Communication*, 20, 1-28.
- Wright, J. (1999). *Idioms Organiser*. Hove, U K: LTP.
- Yorio, C. (1980). Conventionalized language forms and the development of communicative competence. *TESOL Quarterly*, 14, 433-442.