Antonymy

A corpus-based perspective

Steven Jones

London and New York

Antonymy

Antonymy is the technical name used to describe 'opposites', pairs of words such as *poor/rich*, *love/hate* and *male/female*. Antonyms are a ubiquitous part of everyday language, and this book provides a detailed, comprehensive account of the phenomenon.

As well as re-appraising traditional semantic theory and re-evaluating existing categories of antonymy, the book raises wider issues, such as:

- Where do new antonyms come from?
- Which pairs can be regarded as 'good opposites'?
- Why do antonyms tend to favour a particular sequence in text?
- Does word class affect the function of antonymy?

These questions are addressed from a corpus-based perspective, using statistical evidence derived from a 280-million-word corpus of newspaper text. Fresh, empirical statements are made about antonymy, and over 350 authentic examples of how 'opposites' function in the text are recorded and analysed.

This book demonstrates how traditional linguistic theory can be revisited, updated and challenged in the corpus age. It is essential reading for anybody with a particular interest in antonymy or a more general concern about semantics or corpus linguistics.

Steven Jones is a lecturer in English Language and Linguistics at the University of Central Lancashire.

Routledge advances in corpus linguistics
Edited by Anthony McEnery
Lancaster University, UK
and Michael Hoey
Liverpool University, UK

Corpus-based linguistics is a dynamic area of linguistic research. The series aims to reflect the diversity of approaches to the subject, and thus to provide a forum for debate and detailed discussion of the various ways of building, exploiting and theorising about the use of corpora in language studies.

1 **Swearing in English**
Anthony McEnery

2 **Antonymy**
A corpus-based perspective
Steven Jones

3 **Modelling Variation in Spoken and Written English**
David Lee

Antonymy

A corpus-based perspective

Steven Jones

London and New York

First published 2002
by Routledge
11 New Fetter Lane, London EC4P 4EE

Simultaneously published in the USA and Canada
by Routledge
29 West 35th Street, New York, NY 10001

Routledge is an imprint of the Taylor & Francis Group

This edition published in the Taylor & Francis e-Library, 2004.

British Library Cataloguing in Publication Data
A catalogue record for this book is available from the British Library

Library of Congress Cataloging in Publication Data
Jones, Steven, 1974–
Antonymy: a corpus based perspective / Steven Jones.
p. cm.
Includes bibliographical references and index.
1. Antonyms—Data processing.

P325.5.A48 J66 2002
401′.43—dc21

2001058888

ISBN 0-203-16625-6 Master e-book ISBN

ISBN 0-203-26085-6 (Adobe eReader Format)
ISBN 0–415–26374–3 (Print Edition)

To Mum and Dad, my favourite antonymous pair

Contents

Tables

Foreword

Thirty years ago, the series which this book initiates would have been inconceivable. The enormous advances in computer architecture and software that have since made corpus linguistics possible were then only imaginable to the visionary few. Pioneering corpus linguists like Randolph Quirk, M. A. K. Halliday and John Sinclair had to work manually or with primitive computer tools and tiny databases, and their research results found only a small audience. It was a period towered over by Noam Chomsky and his rejection of authentic data as a basis for theory-making made the publication of corpus-based research hard to achieve. Key corpus-based works of that period, such as the OSTI reports by Halliday's research team (Huddleston *et al.* 1968) on scientific English and by Sinclair, Jones and Daley (1970) on collocation, remained unpublished despite their ground-breaking quality, though they were widely circulated in mimeographed form.

Thirty years from now, it is possible that the series that this book initiates will once again seem inconceivable. To a future generation of linguists, natural language processing and corpus-based approaches to linguistic questions may well be so ubiquitous, so normal, that they will think it quaint that a series such as this had ever been needed to provide a home for ground-breaking corpus-linguistic work. Mike Scott, whose WordSmith concordancing software package (Scott 1999) is widely used to analyse corpora, has remarked that talking of corpus linguistics is like talking of 'spade gardening'. His observation has a number of implications. A gardener who never used a spade would be a strange kind of gardener (and a gardener who used nothing but a spade would be even stranger). On the other hand, just as the use of a spade does not constitute a special kind of gardening, so the use of a computer to analyse a corpus as a means of exploring linguistic questions does not constitute *per se* a special branch of linguistics. So why this series and this book?

A clue to the answer to this question comes from an analogy drawn by John Sinclair. He has commented that the computer corpus is to the contemporary linguist what the microscope was to the naturalist of the Enlightenment. The microscope showed the world of nature in a new way, it revealed aspects of nature that had previously only been guessed at (or had been wrongly guessed at) and it opened up opportunities for investigation that in due course

revolutionised the naturalist's world and produced the modern discipline of microbiology. The same impacts are claimed for the computer corpus. 'Microscope biology' is as absurd a label as 'spade gardening' but in the early days of the microscope it was important to explore what it could do and how best to use it. If the computer corpus is our microscope, we need to learn how to use it best.

I rather imagine that the first naturalists to use the microscope rushed around excitedly inviting fellow naturalists to look at the tubular shape of hairs under the microscope or at the veined wings of butterflies. I rather imagine too that some naturalists carried on about the business of observing and classifying in the woods with no other help than a knowledge of the subject and a sharp pair of eyes. I imagine that these field-working naturalists resented the implication that what they were doing was 'old hat' and suspected that what their microscope-obsessed compatriots were doing was rather trivial, while the microscope users for their part knew in their hearts that the future lay with them.

The reason I imagine all this of the early naturalists is that the picture I paint is certainly true of the current linguistic scene. Corpus linguists are so excited by what they do that they are guilty at times of an unjustified contempt for those who use other methods (how dare you use a hoe now we have a spade!). Traditional linguists for their part sometimes resent the way their careful labours are treated as outmoded, and they accuse corpus linguists of theoretical superficiality and of a tendency to focus on the trivial. Corpus linguists *en masse* have certainly on occasion been guilty of a tendency to trivial observation (look at the detail on that butterfly wing!). Not every collocation and word count is of major significance. But despite the sometimes valid criticisms of the traditional linguist, the editors of this series believe that the future lies with those who allow the computer to shape their investigations and to open new avenues of enquiry. They see no conflict between a corpus-based investigation and theoretical rigour, nor between careful attention to the evidence and powerful generalisation. Indeed it is because corpus linguistics and natural language processing work needs to resolve those conflicts, needs to go beyond mere excitement at what can be seen through the microscope, that this series has come into being. The intention is that each volume in the series should either be innovative in the way that it uses the computer and computer corpora in linguistic description or at very least represent current best practice.

There are, however, a number of ways in which a corpus linguistic study may be innovative or an instance of good practice. I referred in the first paragraph to the pioneering corpus linguists of the late 1960s and early 1970s. It is instructive to look at the different ways these linguists worked because in some respects their research anticipates the kinds of study that a series such as this will be publishing. Halliday, working with a distinguished research team including Rodney Huddleston, Richard Hudson and Eugene Winter, took an existing theory, the scale and category grammatical theory that was later to develop into systemic functional grammar, and applied it to a corpus of

scientific writing. Each clause was coded in accordance with his grammatical description, and the results were analysed from a range of perspectives. The objectives were simultaneously to test the theory and to identify the characteristic grammatical features of scientific prose. Both objectives – testing a theory and investigating a well-defined set of data – are echoed in current corpus linguistic work. Thus in future books in this series a theory is tested and either supported or found wanting on the basis of analysis of a corpus. In other books a corpus drawn from a specialist area is investigated and a detailed account is developed of the register(s) and/or genre(s) in question. Occasionally the two objectives may even come together as in Halliday's original work.

John Sinclair and Sue Jones took a different route in their early work. Rather than bringing a fully fledged theory to the corpus they used, they wanted to explore what the computer could reveal about their data. They did not annotate their corpus and brought instead a question to the data: could the computer identify collocations automatically? (The answer seems obvious now, but it was not so then.) This series contains studies that explore different ways of utilising the computer's power to answer questions that were hitherto unanswerable.

Randolph Quirk and his team did not use computers, at least not initially. Instead what they did was create manually an annotated corpus with a view to discovering how the English Language was grammatically organised. Their purpose was to uncover the facts of usage that had hitherto been only anecdotally recorded. Theory was not primary; description was everything. The facts of the language were assumed to be disparate and not necessarily amenable to neat theoretical encapsulation. This series will contain studies that are the result of simply looking 'through the microscope'. The book you are holding is one.

Steven Jones takes the existence of antonymy for granted, at least as a psychological reality, but he recognises that most research into the phenomenon has been introspective and untested against data. Accordingly his concern is to see what a computer corpus can reveal about the manner in which antonymous pairs are used. He looks 'through the microscope' and discovers that antonyms are typically used in ways rather different from those posited in the literature (and popular imagination).

Steven Jones' work was undertaken at the University of Liverpool as part of a research project (ACRONYM) run by Antoinette Renouf and funded by the EPSRC. The main purpose of this project was to develop software capable of automatically identifying the 'nymic' relations of the language (synonyms, antonyms, hyponyms, meronyms, etc.). Amongst the methods used to achieve this end was the exploration of grammatical frameworks that characteristically were associated with 'nyms' and could be used as a discovery procedure to elicit these 'nyms' (Renouf 1996; Collier, Pacey and Renouf 1998; Hearst 1998). Steven Jones' work on the frameworks that can be used to elicit antonyms was the starting point of this book.

There are a number of respects in which Steven Jones' book represents

current best practice and pushes the boundary of corpus linguistics, and his work on the frameworks that can be used to discover future antonyms is one of them. His procedure is a simple one. He takes known antonyms (*good/bad, hot/cold, rich/poor*, etc.) and examines the way they are used when they occur within the same sentence. On the basis of this he then intuitively identifies certain major (and minor) categories of use into which the vast majority of antonymous pairs can be sorted; these categories represent largely distinct functions served by antonymous pairs in the construction of messages. These functions indeed are, Steven Jones argues, the rationale for the language having recognisable antonyms in the first place. The fact that such sorting into functions is unproblematic and leaves only a tiny residue of instances unaccounted for serves as a verification of the original intuited categories. These categories are in turn examined for grammatical regularities ('frameworks') which have as their apparent function the signalling of antonymous pairs. Finally a handful of the frameworks are then tested for their ability to elicit antonymous pairs and some intuitively convincing 'new' antonyms are found.

The finding that 'most functions of antonymy are closely associated with certain lexico-syntactic frameworks' (p. 177) confirms the validity of the positions of Renouf (1996) and Hearst (1998) and is a significant finding in its own right. The use of frameworks to identify antonyms in the making offers an exciting way forward for students of lexical relations, whether or not they endorse the categories that Jones establishes. Jones himself is clear about the importance of intuition in the study: 'the process of creating and classifying a suitable database does not become an exact science simply because a corpus-based approach is taken. Corpus linguistics is not intuition-free linguistics' (p. 44). Importantly, though, the intuition comes in only once – in the analysis of the data – and not twice as would be the case in the more traditional study – in the invention of the data as well as the analysis. As Jones again puts it, 'corpus data helps to eliminate the element of chance by tapping into not one mental lexicon but thousands' (p. 21).

A second respect in which Jones' work represents current best practice is in his attitude to the awkward examples in his data. Traditional linguistic practice has placed great emphasis on the importance of the counter-example, the sentence that will not accommodate itself to the theoretical statement meant to account for it. Such sentences were almost inevitably fabricated, which did not of course affect their validity except where their grammaticality was questioned, but they were invented as part of an argument and the inventor usually had an alternative theoretical explanation in waiting. In corpus linguistics, the situation is quite different. Counter-examples come unbidden as part of no argument and the availability of sufficient explanations can certainly not be taken for granted. Perhaps, though, corpus linguistics has not taken the importance of counter-examples sufficiently seriously; the overwhelming tendency has been to talk of overwhelming tendencies. This is not however the strategy adopted by Jones. He is not content with merely illustrating his categories. Every sentence has to be either accounted for or

explicitly stated not to have been; where his categories leak, they are shown to leak. Paradoxically this ought to give us confidence in those categories, because the counter-examples come from those thousands of mental lexicons already referred to and not from the single mental lexicon of a linguist with a point to demonstrate.

The final respect in which Steven Jones' work is important lies in his novel stance towards collocation – and it is perhaps this strength that is most easily overlooked because it pervades the whole book. The standard take on collocation is that it can be identified on the basis of statistical information with the help of large corpora. One starts with the 'seed' word and this is used to generate a list of collocations from the most central to the most marginal. The nature of these collocations is sometimes examined from the point of view of the grammatical relations between the words (Hunston and Francis 1998; Francis, Hunston and Manning 1996, 1998) or from the point of view of the semantic relation that underpins it (Sinclair 1991; Louw 1993; Stubbs 1996; Hoey 1998). What is rarely looked at is the *function* of the collocation.

Jones however does just this. He assumes the collocates as a cultural given. Thus he takes it as his starting-point that *old* collocates with *new*, that *rich* collocates with *poor*, that *right* collocates with *wrong*. This is not to say that his evidence for claiming their collocation is not strong – he rejects, for example, those antonymous pairs posited by Roget that have no support in his corpus – but it is to say that he is not interested in demonstrating that antonyms are collocates. (We should be grateful for this; a monograph that told us that *man* collocates with *woman* would not be exciting.) What he is instead interested in is what the antonymous collocates are used for. In other words, he takes the argument about collocates one stage further. There are many studies of collocation but rather fewer that focus on the functions of particular collocations. Jones demonstrates that these collocations have regular functions for the message as a whole. It remains to be seen whether other types of collocation can be shown to have regular and categorisable functions, but at the very least it would be reasonable to see whether similar claims can be made about synonymous, meronymous and hyponymous collocations. If it proves to be the case that other kinds of collocations are indeed functioning in regular and interesting ways in the way that Jones shows that antonymous pairs do, then he will have opened up a new area of inquiry for corpus linguists. If on the other hand it proves to be the case that antonyms are alone in having this characteristic, it will support Jones' claim that antonyms have a unique psychological reality for users of the language and are therefore especially worthy of study. Either way this book serves as a valuable contribution to work in corpus linguistics and as a worthy beginning to a series dedicated to the use of one of the most valuable tools in the linguist's tool shed. And if in thirty years' time the use of a corpus is as obvious as the use of a spade in gardening or the use of a microscope in nature study, then we will all be well pleased.

Michael Hoey
University of Liverpool

References

Collier, A., Pacey, M. and Renouf, A. (1998) 'Refining the Automatic Identification of Conceptual Relations in Large-scale Corpora', in E. Charniak (ed.) *Proceedings of the Sixth Workshop on Very Large Corpora*, Montreal: University of Montreal.

Francis, G., Hunston, S. and Manning, E. (eds) (1996) *Grammar Patterns: 1. Verbs*, Collins Cobuild, London: HarperCollins.

Francis, G., Hunston, S. and Manning, E. (eds) (1998) *Grammar Patterns: 1. Nouns and Adjectives*, Collins Cobuild, London: HarperCollins.

Hearst, M. A. (1998) 'Automated Discovery of WordNet Relations', in C. Fellbaum (ed.) *WordNet*, Cambridge, Mass.: MIT Press.

Hoey, Michael (1998) 'From Concordance to Text Structure: New Uses for Computer Corpora', in *Proceedings of PALC '97*, Lodz: University of Lodz.

Huddleston, R., Hudson, R., Henrici, A. and Winter, E. O. (1968) 'Sentence and Clause in Scientific English: Report to OSTI on the Research Project "The Linguistic Properties of Scientific English"', London: University College London, Department of General Linguistics.

Hunston, S. and Francis, G. (1998) 'Verbs Observed: A Corpus-driven Pedagogic Grammar of English', *Applied Linguistics*, 19: 45–72.

Louw, Bill (1993) 'Irony in the Text or Insincerity in the Writer? The Diagnostic Potential of Semantic Prosodies', in M. Baker, G. Francis and E. Tognini-Bonelli (eds) *Text and Technology: In Honour of John Sinclair*, Amsterdam: John Benjamins, pp. 157–76.

Renouf, Antoinette (1996) 'The ACRONYM Project: Discovering the Textual Thesaurus', in J. M. Aarts, P. De Haan and N. H. J. Oostdijk (eds) *Synchronic Corpus Linguistics – Papers from the Sixteenth International Conference on English Language Research on Computerised Corpora*, Amsterdam: RoDoPi.

Scott, Mike (1999) *WordSmith Tools, Version 3*, Oxford: Oxford University Press.

Sinclair, J. McH., Jones, S. and Daley, R. (1970) 'English Lexical Studies: Report to OSTI on Project C/LP/08', Birmingham: University of Birmingham.

Sinclair, J. McH. (1991) *Corpus, Concordance, Collocation*, Oxford: Oxford University Press.

Stubbs, Michael (1996) *Text and Corpus Analysis: Computer Assisted Studies of Language and Culture*, Oxford: Blackwell.

Acknowledgements

As co-editor of this series, I am indebted to Anthony McEnery for practical guidance and astute criticism. As co-supervisor of the doctoral work upon which this book is based, I am indebted to Antoinette Renouf for warm friendship and insightful professional advice. At different times, my research has benefited immeasurably from your experience, expertise and generosity.

As both co-editor of this series and co-supervisor of the doctoral work upon which this book is based, I am indebted to Michael Hoey for remarkable academic tutelage. As an undergraduate, your passion for English Language intrigued me; since then, your measured, convivial advice and insatiable intellectual hunger has been a major source of inspiration.

On a more general note, I am grateful to innumerable colleagues, former colleagues, students and former students at the following institutions: the University of Liverpool; the University of Central England; and the University of Central Lancashire. My PhD research was undertaken at the *Research and Development Unit for English Studies* and funded by the EPSRC as part of the ACRONYM project.

On an individual level, I would like to thank Alex Collier, Brian Francis, Mike Pacey and Phill Ralph (for computational and statistical assistance); Kara Springer (for daily electronic support); Howard Jackson, Mike McCarthy and Mike Scott (for crack academic input); Vikki Hulse (for masterful proofreading); John Owens (for arsey feedback); Eithne Quinn (for serious thinking and boss ideas); my family and friends in Liverpool (for much needed escapism); Woody Allen and Everton Football Club (for laughter); and Victoria Morgan (for all of the above, except the computational and statistical assistance, and to whom this book owes most).

1 The 'unique fascination' of antonymy

The title of this chapter is borrowed from an expression used by Cruse (1986: 197) when describing what makes antonymy special. Though often grouped together with synonymy, hyponymy and meronymy, the scope of antonymy is much greater than that of its fellow sense relations. For proof of this, talk to my four-year-old nephew, Thomas. He understands the concept of 'opposites' and excitedly tells me all about pairs such as *big*/*little*, *boy*/*girl* and *happy*/*sad*. Together with other childhood learning exercises (such as counting, reciting nursery rhymes and distinguishing between colours), recognising antonyms seems to be a natural stage in an infant's linguistic development. This is not something which could be said of other sense relations. Furthermore, our exposure to antonyms is not restricted to childhood; we are surrounded by 'opposites' throughout our adult life and encounter them on a daily basis.[1] This book will begin by examining the prominence of antonymy in contemporary society and discuss why it seems to have transcended its role as a mere sense relation.

'Opposites' or 'antonyms'?

Though Thomas can happily recall an almost endless list of 'opposites',[2] he is not familiar with the term 'antonymy'. Often, 'antonymy' is thought of as being the correct linguistic term for 'opposites', as 'parentheses' is simply a technical name for 'brackets'. However, this is not entirely true. Some linguists, such as Lyons (1977) and Cruse (1986), apply the label of 'antonymy' to pairs such as *heavy*/*light*, *new*/*old* and *fast*/*slow*, but do not accept that pairs such as *alive*/*dead*, *false*/*true* and *female*/*male* are antonymous.[3] This creates a problem as both sets of pairs would be readily identified as 'opposites' by any native speaker of English. I propose to resolve this problem by using the term 'antonymy' in its broader sense, referring to any pair of words which could be intuitively recognised as 'opposites'.

This semantic tension between 'antonyms' and 'opposites' may partly account for the tendency of commentators to shy away from recognising the status held by the phenomenon in language. For example, Lyons is wary of the word 'antonymy' because it is 'hardly more precise in the usage of most

authors than the word *oppositeness* which it replaces' (1977: 270) and Simpson is equally cautious about what he refers to as a 'catch-all category' (1997: 72). But this instinct to reject a general term needs to be examined. Whilst it is true that antonymy encompasses a multitude of relationships, each slightly different from the next, it is equally true that all established 'opposites' in English share something in common. Any native speaker would immediately identify the 'opposite' of words such as *cold*, *legal* and *above* without feeling the need to distinguish between gradable antonymy, complementarity and converseness. To deny the status of antonymy to any familiar pair of 'opposites' seems counter-intuitive and likely to obscure the underlying uniformity of all such word pairs.

Why study antonymy?

The simple, though somewhat clichéd, answer to the question of why antonymy should be studied is 'because it's there'. Indeed, I would suggest that antonymy is much more 'there' than many other linguistic phenomena which have received considerably more attention in recent times. It is the 'most readily apprehended' (Cruse 1986: 197) of sense relations and many examples of antonymy become deeply ingrained in our mental lexicon from infancy. 'Opposites' are quickly learnt and rarely forgotten. Furthermore, it has even been suggested that antonymy has a magical quality, as Cruse explains:

> Opposites possess a unique fascination, and exhibit properties which may appear paradoxical. Take, for instance, the simultaneous closeness, and distance from one another, of opposites. The meanings of a pair of opposites are felt intuitively to be maximally separated. Indeed, there is a widespread idea that the power of uniting or reconciling opposites is a magical one, an attribute of the Deity, or a property of states of mind brought about by profound meditation, and so on . . . Philosophers and others from Heraclitus to Jung have noted the tendency of things to slip into their opposite states; and many have remarked on the thin dividing line between love and hate, genius and madness, etc.
>
> (Cruse 1986: 197)

Cruse's observations echo earlier rhetoric about the significance of antonymy. For instance, despite rejecting 'antonymy' as a useful label, Lyons, notes 'a general human tendency to categorize experience in terms of dichotomous contrast' (1977: 277). Such claims are difficult to substantiate but will now be examined in relation to a number of fields in which antonymy has been said (or can be said) to play an important role: first, research into the acquisition of language in children will be considered to determine why antonyms are memorised in infancy; second, the psychological domain of word association testing will be explored to determine whether the human mind can be said to

operate in terms of oppositions; third, the role of antonymy in literature will be analysed to speculate why 'opposites' often feature heavily in dramatic or memorable prose; and fourth, metaphor will be examined with a view to illustrating the extensive role served by antonymy in figurative speech. Collectively, the analyses of these four discrete areas may help to illuminate the extent to which antonymy can be seen as a fundamental experience-organising mechanism.

Language acquisition

It has been widely documented that children tend to grasp the concept of oppositeness at a very early age, often learning antonyms in pairs rather than as single items. For example, Kagan notes that 'soon after learning the meaning of *up*, the child learns the meaning of *down*; after learning the meaning of *high*, she learns *low*; after *good*, she develops the meaning of *bad*' (1984: 187). This could reflect the tendency to dichotomise which Lyons notes; alternatively, it could simply be a learning strategy used by children as part of their general language acquisition mechanism. It seems efficient to learn closely related words in tandem, yet it is difficult to think of other word pairs which are learnt in the same fashion as antonyms. One would not necessarily feel a similar urge to learn synonyms in unison, nor would one find it problematic to fully understand a superordinate term without first being taught all of its corresponding hyponyms. However, in each of Kagan's examples, it is difficult to conceptualise one antonym without first having some notion of the other. Can one fully comprehend the meaning of *up*, *high* and *good* without having any concept of *down*, *low* and *bad*?

Conducting research into second language acquisition, Miller and Fellbaum note that 'when given only one member of an antonymous or opposed verb pair, [students of a foreign language] will insist on being taught the other member'; likewise, when referring to native language competence, Egan states that 'it is good, we feel, to know the exact antonym of a word, for not only will it give us a firmer grasp of the meaning of the word to which it is opposed, but inversely, of itself' (both cited by Muehleisen, 1997: 4). These observations suggest that adults favour antonymy as strongly as children – we are drawn to 'opposites' when learning a new language and feel more comfortable with the precise meaning of a word in our native tongue if we are familiar with its corresponding antonym.

Whether this gives antonymy a 'unique fascination' is difficult to judge. Indeed, the integral position held by antonymy in the mental lexicon is something of a chicken-and-egg situation. It is almost impossible to know whether language simply reflects existing oppositions in the outside world or whether we, as humans, are subconsciously predisposed to impose such dichotomies. Whatever the cause, the consequence is that 'opposites' hold a key place in language and this is reflected by the pull of antonymy to learners of a first or other language.

Associationism

Another field in which antonymy has been identified as playing a significant role is that of word association testing. For example, Clark (1970) examined the tendency shown by informants to provide the antonym of a stimulus word when asked to 'say the first thing that comes into your head'. He concluded that:

> If a stimulus has a common 'opposite' (an antonym), it will always elicit that opposite more often than anything else. These responses are the most frequent found anywhere in word association.
>
> (Clark 1970: 275)

Clark's conclusions follow those of Deese (1964), who found forty words all of which elicited their antonym most commonly.[4] Indeed, it is difficult to dispute the fact that people often think in terms of oppositions when faced with a word association test. However, data shows that such tests also elicit synonyms and general collocates (Clark 1970: 281–2); antonymy is only one of many relationships reflected by informants. Indeed, the results of word association tests, though often cited as being indicative of the central role of antonymy in the mental lexicon, are not particularly convincing. Such is the artificiality of the procedure, I suspect that responses may disclose no more about cognitive practices than they allegedly do about an informant's sexual hang-ups (Jung 1973: 288–317).

Furthermore, this research should be examined within the context of the dominant linguistic schools of the day. Clark is puzzled by his findings because 'language, the critics say, should not be thought of as a consequence of built-up associations' (1970: 272). The 'critics' referred to here would include Chomsky (1965) and other proponents of transformational grammar theories. However, many contemporary linguists would be much more receptive to seeing language exactly as here described (and dismissed) – as a series of 'built-up associations'. Therefore, though Clark's findings remain of interest, their repercussions are perhaps less significant now than they were in the mid-1960s.

Perhaps the most interesting aspect of Clark's work is the phrase he coins to describe antonym elicitation: 'the minimum contrast rule' (1970: 275). Intuitively, one might not feel that antonyms reflect 'minimum contrast'. Indeed, one might be more inclined to think of antonyms as having maximum contrast. Yet Clark's thinking is quite revealing – the contrast, he believes, is minimum because antonyms only differ in one respect. For example, *girl* elicits *boy* because they are both human and both non-adult. They only differ against one scale, namely the scale of gender. This thinking, which closely resembles what would later be termed componential analysis (Leech 1974: 89), sheds light on Cruse's remark about 'the simultaneous closeness, and distance from one another, of opposites' (1986: 197). Cruse does not refer to the minimum contrast rule directly, but is clearly describing the same phenomenon as Clark when he writes:

> This paradox of simultaneous difference and similarity is partly resolved by the fact that opposites typically differ along only one dimension of meaning: in respect of all other features they are identical, hence their semantic closeness; along the dimension of difference, they occupy opposing poles, hence the feeling of difference.
>
> (Cruse 1986: 197)

The idea that a pair of antonyms are semantically alike may initially seem counter-intuitive, but it is difficult to fault Cruse's logic. On the semantic scale along which they operate, antonyms are some distance apart; however, in many other respects (word class, paradigmatic interchangeability, collocational profile, etc.) antonyms are remarkably similar.

Literature

> It was the best of times, it was the worst of times, it was the age of wisdom, it was the age of foolishness, it was the epoch of belief, it was the epoch of incredulity, it was the season of Light, it was the season of Darkness, it was the spring of hope, it was the winter of despair, we had everything before us, we had nothing before us, we were all going direct to Heaven, we were all going direct the other way – in short, the period was so far like the present period, that some of its noisiest authorities insisted on its being received, for good or for evil, in the superlative degree of comparison only.
>
> (Charles Dickens, *A Tale Of Two Cities*)

The opening few lines of *A Tale Of Two Cities* are among the most familiar ever written by Charles Dickens. They are also replete with antonyms. In total, nine contrasts are presented in the text above, most of which are established antonymous pairs. The sentence (for the above is all a single orthographic sentence) features no coordinators to link its clauses, but rather relies on apposition. The effect of the antonymy is to describe a period of confusion, contradiction and extremity. The text has an encompassing quality; the repetition of antonymous pair after antonymous pair evokes a feeling of being told much about the time, yet being told nothing because each new clause is effectively cancelled out by the next. One could interpret the passage as a dialogue which reflects two conflicting opinions about the period, or one could attribute the narrative to a single voice. Either way, given the magnitude of Dickens's work, it is curious that these lines are perhaps the best known of all. Curious, and possibly reflective of 'the unique fascination' (Cruse 1986: 197) of antonymy.

Another example of a memorable and important piece of literature making extensive use of antonymy can be found at the beginning of the Bible:

> And God said, 'Let there be light,' and there was light. God saw that the

> light was good, and he separated the light from the darkness. God called the light 'day' and the darkness he called 'night'. And there was evening, and there was morning – the first day.
>
> (Genesis 1: 4–5)

The passage above is part of a series of dichotomous distinctions made in the first chapter of Genesis, where Creation is explained largely in terms of antonymy. Even the Bible, an ancient text written in a Semitic language, uses 'opposites' extensively. Indeed, in the field of theological deconstruction, Taylor states that 'Christian theology is repeatedly inscribed in binary terms' (1984: 8) and it could be argued that many religions are built around a system of antonymous extremities such as *God* and the *Devil*, *Heaven* and *Hell*, *good* and *evil*, etc.

Metaphor

Further proof that antonymy is 'beneath the skin' of language can be found if the role of opposition within metaphor is examined. So deep-rooted are antonymous concepts in language, several pairs have moved beyond their literal meaning and function figuratively in many contexts. In their seminal work on the pervasiveness of metaphor in language, Lakoff and Johnson (1980: 14) identify a number of orientational metaphors which show concepts being systematically expressed according to directional terms. Below are some examples of the antonymous pair *up*/*down* acting metaphorically, taken from Lakoff and Johnson (1980):

HAPPY IS UP; SAD IS DOWN	my spirits *rose*; I'm *depressed*
CONSCIOUS IS UP; UNCONSCIOUS IS DOWN	wake *up*; he *fell* asleep
MORE IS UP; LESS IS DOWN	my income *rose*/*fell* last year
VIRTUE IS UP; DEPRAVITY IS DOWN	she has *high* standards; that was a *low* trick

Dozens of examples similar to those recorded above are cited by Lakoff and Johnson, and it is clear that antonymous relations are frequently exploited in the metaphoric system which underpins English. Indeed, it is not difficult to think of other ways in which antonymous pairs are used metaphorically in British culture: for example, political ideology is characterised in terms of *left* or *right* so extensively that phrases related to this antonymous scale (*left-of-centre*, *hard-right*, etc.) are often used to characterise the beliefs of individual politicians.

The idea of language being structured in terms of antonymous distinctions is taken even further by Cixous. Working within the area of feminist literary theory, she argues that 'thought has always worked through opposition' (1989: 91), claiming that linguistic dichotomies reflect/create disequalities between the sexes. Though some might find this position extreme, the fact

that Cixous believes that 'opposites' are at the core of human thought adds weight to the claim that antonymy is a crucial and under-analysed linguistic phenomenon.

Omnipresent antonyms

This chapter has sought to prove that antonymy is a key feature of everyday life. Should further evidence be required, try visiting a public lavatory without checking which is the 'gents' and which is the 'ladies'. On your way out, ignore the instructions which tell you whether to 'push' or 'pull' the door. And once outside, pay no attention to whether traffic lights are telling you to 'stop' or 'go'.[5] At best, you will end up looking very foolish; at worst, you will end up dead.

Antonymy holds a place in society which other sense relations simply do not occupy. Whether or not there exists a 'general human tendency to categorize experience in terms of dichotomous contrast' (Lyons 1977: 277) is not easily gauged, but, either way, our exposure to antonymy is immeasurable: we memorise 'opposites' in childhood, encounter them throughout our daily lives, and possibly even use antonymy as a cognitive device to organise human experience.

Does this give antonymy a 'unique fascination'? That question cannot be answered objectively; however, we are now in a position to speculate further about why antonymy occupies such a central place in the mind. What Cruse refers to as the 'paradox' of antonymy can be explained by the fact that antonyms are effectively a special kind of co-hyponym. For example, *female* and *male* are both adjectives (or nouns) which define gender; *hate* and *love* are both human emotions; *bad* and *good* are both quality-measuring attributes of a given concept, and so on. The word which has a maximum opposition with *happy* is not *unhappy* or *sad* (for they are both adjectives and they both describe one's feelings); rather it would be a word such as *cutlery*,[6] which shares nothing in common with *happy*. By definition, antonyms have lots in common, a fact which underpins Clark's formation of the minimum contrast rule.

This commonality also helps to account for another of Cruse's observation, namely 'the frequency of speech errors in which the intended word is substituted by its opposite' (1986: 197). This is illustrated by the quote below, taken from the *Liverpool Echo* (16 August 1999):

> These chants show a complete lack of disrespect to a lad who is now admired and has a child, and those supporters should put themselves in his shoes.

It would appear that *disrespect* has been inadvertently substituted for *respect* in this context. This is perhaps symptomatic of the psycholinguistic closeness of antonyms; such slips of the tongue are not immediately evident and this one appears to have escaped the attention of all newspaper staff who reported and

published it. Indeed, even American presidents seem susceptible to errors of antonym substitution, as the following blunders show:

> We're trying to get unemployment to go up and I think we'll succeed.
> (Ronald Reagan, 18 October 1982)

> I do not believe we have put a guilty – I mean an innocent – person to death in Texas.
> (George W. Bush, 16 June 2000)

These speech errors further demonstrate the similarity between antonyms, but this similarity hardly accounts for 'the thin dividing line between love and hate, genius and madness, etc.' (Cruse 1986: 197), nor does it corroborate any suggestion that magical powers are required to unite opposites. At best, it is a linguistic reflection of the myths which surround antonymy. However, if we are to ask why such myths have evolved, the answer is surely that 'opposites' are more widespread and more primal than other relations holding between words. For this reason, antonymy is examined afresh in this book, with corpus data being used extensively to help shed new light on an enigmatic (yet fundamental) linguistic phenomenon.

2 A brief history of antonymy

Given that antonymy is 'the only sense relation to receive direct lexical recognition in everyday language' (Cruse 2000: 167), one might expect to find that a robust system of classification has emerged. However, though antonymy has been the subject of continual linguistic scrutiny, not all analysts have reached the same conclusions. Granted, most commentators categorise antonymous pairs according to broadly similar criteria, but the terminology used to describe these categories is anything but standard. Furthermore, an adequate definition of antonymy has yet to be agreed upon. Though all speakers can easily recall a lengthy list of 'opposites', describing this relation in a clear, concise fashion has proved problematic. Antonymy is a phenomenon better suited to exemplification than definition.

Defining antonymy

The word 'antonymy' was coined in 1867 by C. J. Smith[1] as an opposite of (and by analogy with) 'synonymy'. Whether antonyms are really the 'opposite' of synonyms is a moot point, especially as it is widely accepted that true synonymy does not exist (e.g. Palmer 1976: 94), but this does not challenge the validity of the concept. Since 1867, numerous attempts have been made to pin down the meaning of antonymy and formulate a workable definition of the term, but the problem is that antonymy lends itself more to illustration than description. Good 'opposites' are intuitively available to us all (*old*/*young*, *down*/*up*, *lose*/*win*, *bad*/*good*, etc.) but finding a definition which adequately accounts for every example of antonymy is more problematic. In general, two ways of defining antonymy have emerged: the first involves semantic criteria; the second involves lexical criteria. The relative merits of each approach will now be evaluated.

Semantic definitions

The *Longman Dictionary of Applied Linguistics* (Richards, Platt and Weber 1985) defines an antonym as 'a word which is opposite in meaning to another word' (1985: 14). Similarly, in his *Dictionary of Linguistics and Phonetics*,

Crystal[2] says of antonymy that 'in its most general sense, it refers collectively to all types of semantic oppositeness' (1985: 18). These definitions both reflect semantic criteria; indeed, antonymy is often defined simply as 'oppositeness of meaning' (see Palmer, who relies on this gloss despite claiming that antonymy 'can be defined fairly precisely' (1976: 94)).

The problem with an exclusively semantic definition of antonymy is that it fails to explain, or even acknowledge, the tendency for certain words to become enshrined as 'opposites' in language while others do not. Antonyms are semantically opposed words, but not all semantically opposed words are antonyms. Consider *rich* and *poor*, a pair of words which operate along a given scale, namely the scale of wealth. *Affluent* and *broke* also operate along this scale, but, intuitively, one would be reluctant to describe them as antonyms (or, at least, one would not want to describe them as having the same degree of antonymity as *rich*/*poor*). Cruse exemplifies this with the words *tubby* and *emaciated* (1986: 262) – without question, these words are 'opposite in meaning' to one another, yet it is difficult to imagine any native speaker of English volunteering them as 'opposites'. Indeed, almost all established antonyms have synonyms which resist the label of antonymy. For example, *heavy* and *light* are better 'opposites' than *weighty* and *insubstantial*; *fast* and *slow* are better 'opposites' than *speedy* and *sluggish*; *happy* and *sad* are better 'opposites' than *ecstatic* and *miserable*, and so on.

One could counter this argument by noting that the examples above feature weak synonyms: *sad* and *miserable* may be similar in meaning to one another, but they are not identical; neither are *happy* and *ecstatic* entirely interchangeable. Though this is true, *miserable* and *ecstatic* still belong firmly to the same semantic scale, that of happiness, and they each belong to opposite halves of this scale. Yet these words lack the antonymous power of *happy* and *sad*.

Therefore, just as individual antonymous pairs operate along a scale, so too does the phenomenon of antonymy. Those pairs towards the top of the scale are readily identified as 'opposites' by native speakers (*cold*/*hot*, *short*/*tall*, *narrow*/*wide*, etc.); those lower down the scale may still reflect contrast and may still be identified as 'opposites', but seem to lack the 'clang' (Miller 1990, cited in Muehleisen 1997: 4) of antonymy (*relaxed*/*tense*, *smart*/*stupid*, *refined*/*uncouth*, etc.); those at the bottom of the scale still express opposite ends of a semantic range, but would almost never be identified as 'opposites' (*cerebral*/*dozy*, *beatific*/*crap*, *gossipy*/*taciturn*, etc.). As Justeson and Katz put it, 'adjectives may be more or less antonymous rather than simply antonymous or not antonymous' (1991: 147). The upshot of this is that our intuitions about 'opposites' cannot be explained purely by reference to semantics; such definitions of antonymy are inadequate.

Lexical definitions

Justeson and Katz resolve the problem of definition by referring to antonymy as a lexical relation, 'specific to words rather than concepts' (1991: 138). They

cite the scale of size as evidence, arguing that although *small* and *little* are synonyms and *big* and *large* are also synonyms, most native speakers of English would be intuitively aware that the more appropriate antonymous pairs are *large*/*small* and *big*/*little*. Semantically, *large* remains diametrically opposed to *little*, but the words themselves are not considered antonymous.[3] These four words are examined in detail by Muehleisen (1997), who finds that neither *big* and *large* nor *small* and *little* are true synonyms because they share different collocational profiles.[4] She concludes that:

> Although *big* and *large*, and likewise *little* and *small*, are near synonyms, this does not mean they are identical in meaning; they are synonyms by virtue of the fact that they are associated with the same semantic dimension, but they are differentiated by the fact that they modify different kinds of nouns . . . Good opposites are adjectives that not only occupy opposite ends of a shared semantic dimension, but also describe the same kind of things.
>
> (Muehleisen 1997: 113)

This reiterates the view that antonymy should be seen as a relationship between words, not a relationship between concepts. One caveat to this is provided by Fellbaum (1995), who points out that antonymy not only arises between pairs belonging to the same word class, but also arises across word class. Thus, writers will find ways of opposing *loving*/*hate*, *love*/*hates*, *loved*/*hatred*, etc. as well as grammatically compatible pairs such as *loves*/*hates* and *loved*/*hated*.

However, any definition of antonymy must be lexical as well as semantic. Antonyms need to have 'oppositeness of meaning' (Jackson 1988: 75), but they also need to have a strong, well-established lexical relationship with one another. Those word pairs which meet both criteria are known as 'prototypical' or 'canonical' antonyms; those word pairs which meet the first criterion but not the second have been dubbed 'peripheral' or 'non-canonical' (terminology provided by Cruse (1986: 198) and Murphy (1994: 4) respectively). These labels essentially refer to those pairs which are lexically enshrined (e.g. *hard*/*soft*) and those pairs which are not (e.g. *malleable*/*rigid*). Inevitably, the more antonymity a word pair is thought to have, the more linguistic attention it has received; currently favoured categories of antonymy tend to be based on prototypical antonyms only.

Traditional categories of antonymy

The meanings of antonymous pairs have been logically dissected by a number of linguists (e.g. Leech 1974; Lyons 1977; Cruse 1986) and antonyms have been categorised according to their theoretical differences, perhaps at the expense of their intuitive similarity. Each of the traditional categories of antonymy will now be explained and exemplified, beginning with the two

major classes of antonymy, which I shall refer to as 'non-gradable' and 'gradable'.

Non-gradable antonymy

Non-gradable antonymy is the name given to antonymous pairs such as *man/woman*, *alive/dead* and *married/unmarried*. Leech refers to this class as 'binary taxonomy' (1974: 109), but the majority of writers (see Palmer 1976; Carter 1987; Jackson 1988) prefer to speak of 'complementarity'. Kempson – whose favoured term is 'simple binary opposition' – describes examples of non-gradable pairs as 'the true antonyms' (1977: 84). However, this description is particularly confusing in light of the unwillingness of other linguists, namely Lyons (1977) and Cruse (1986), to acknowledge complementarity as a form of antonymy at all. Their feeling is that the label 'antonymy' should be restricted to gradable pairs.

The criterion necessary for an opposition to be considered non-gradable is that the application of one antonym must logically preclude the application of the other. For instance, if X is a *man*, X cannot be also a *woman*; if X is *married*, X cannot also be *unmarried*, and so on. In theory, one could contest this criterion by claiming that, for example, *motorbikes* are neither *men* nor *married*; does this make them *unmarried women*? To counter such purely hypothetical objections as these, Leech turns to the area of componential analysis (1974: 98). This enables the words *man* and *woman* to be expressed as follows:

man:	+ HUMAN	+ ADULT	+ MALE
woman:	+ HUMAN	+ ADULT	− MALE

Thus, *man* and *woman* are regarded as antonyms because of the clash between the semantic features +MALE and –MALE. This opposition is licensed by the congruence of +HUMAN and +HUMAN (which eliminates the motorbike hypothesis) and of +ADULT and +ADULT (which acknowledges that *girl* could be seen as the antonym of *woman* if the gender scale were equal but the age scale different). In other words, inherent to the words *man* and *woman* are the notions of humanness and adulthood.

Leech (1974) and Kempson (1977) both draw heavily on componential analysis in their analyses of antonymy and this strategy is effective when dealing with certain antonymous pairs, especially those which concern kinship terms or gender. However, the explanatory power of componential analysis does not seem to extend beyond this – describing an antonymous pair such as *bachelor/spinster* is unproblematic, but tackling a pair such as *active/passive* creates many more difficulties. To begin with, the fact that these words are not nouns means that they resist neat superordinate labels such as 'adult'. This particular antonymous pair could be applied to animate nouns (*active participant*; *passive majority*) or to inanimate nouns (*active trading*; *passive smoking*). To apply a componential analysis, one would struggle to define the

pair as being anything more specific than +DESCRIPTIVE. Whereas *active* is +DOING SOMETHING, *passive* is –DOING SOMETHING. The problems here is that whereas +HUMAN +ADULT +MALE clearly defines *man*, +DESCRIPTIVE –DOING SOMETHING defines *passive* with significantly less precision.[5] Therefore, though a useful tool for describing some antonymous pairs, componential analysis is not always, as Kempson believes, 'a much more explicit, clear and economical way of characterising . . . relations' (1977: 86).

At a fundamental level, Jackson questions the entire concept of non-gradable opposition, claiming that 'just about any non-gradable antonym may be made gradable' (1988: 76). He challenges the *female*/*male* opposition on the grounds of sex-change operations and chromosome research, and Cruse is similarly troubled by the effect that zombies and vampires have on the *alive*/*dead* distinction (2000: 169). However, a challenge to the concept of non-gradable antonymy may also be lodged using corpus evidence, as the following sentences[6] illustrate:

1a But I feel much **more alive** when I'm acting – the rest of life becomes much more interesting.
1b Josh Logan had noted he was all a director could hope for: tall, humorous, **extremely male**.
1c Margo at that time was **very pregnant** with Hector, and we had dinner and talked late into the night.

Examples such as those above remain relatively rare, but hardly strike the reader as being incomprehensible. Technically, one is either *alive* or *dead*, but that is no impediment to our understanding of how the writer of sentence 1a is feeling. Similarly, to describe someone as being *extremely male* or *very pregnant* may attract censure in some quarters (newspapers' letter pages often reflect readers' distaste for 'very unique'), but we are well aware what is meant by these phrases: an *extremely male* person is macho in some respect; a *very pregnant* woman is simply in the latter stages of pregnancy. Lyons notes that 'what we are grading, presumably, are various secondary implications, or connotations' (1977: 278). This presumption is fair, but Lyons goes on to state that 'recognition of the possibility of grading normally ungradable antonyms . . . does not imply that there is not a sharp distinction to be drawn between gradable and ungradable antonyms in a language-system' (1977: 279). In strictly logical terms, Lyons may be right – the division between gradable and non-gradable antonymy is fundamental in theory – but it is difficult to understand why this distinction should be considered so important if textual evidence, as exemplified by the three sentences above, continually fudges the line between the two categories in practice. Again, corpus evidence can be usefully invoked in this issue and a statistical analysis of whether non-gradable antonyms function differently in text from gradable antonyms is presented in Table 9.3 (p. 150).

When it comes to intuitive categories of antonymy, the system developed

by Cruse (1986) is the most comprehensive and the most complex. In general, Cruse adopts Lyons' terminology but he also creates further sub-categories. In particular, Cruse formulates the notion of the 'antonymous triplet' (1986: 201), arguing that antonymous verbs first need a given condition to be met in order for their contrastive potential to be realised.

For example, if the binary pair *live/die* is to hold true, some instance of *birth* must first occur. Thus, Cruse creates the triplet *be born:live:die*, labelling it 'reversive complementarity'. This is because the outer pair (*be born* and *die*) are reversives, i.e. they denote change in opposing directions. Similar examples include *learn:remember:forget* and *arrive:stay:leave.* The second kind of triplets identified by Cruse are known as 'interactives'. These pertain to antonymous pairs which are themselves a response to a given stimulus. For example, the complementaries *obey* and *disobey* are binary responses to a given command. *Command* is therefore said to be 'interactive' with both *obey* and *disobey*. Other 'interactive' triplets include *request:grant:refuse*, *greet:acknowledge:snub* and *tempt:yield:resist*. The term 'satisfactive' is given to the relationship between a word signifying an attempt to do something and its corresponding success-measuring antonyms. *Compete:win:lose* is an example, as are *try:succeed:fail* and *aim:hit:miss*. Cruse's final category of complementary triplet involves 'counteractives'. Here, the first term of the triplet represents an aggressive action, the second a measure to neutralise it, and the third a consequence of the inability to neutralise it. *Attack:defend:submit* is the primary example, though *charge: refute:admit* and, in the context of a football match, *shoot:save:let in* are other feasible illustrations.

Cruse acknowledges that the final two categories are less fundamental than the first two and I would add that a number of his illustrations of non-gradable antonymy are questionable (*not winning* does not inherently entail *losing*; a valid third outcome is often still possible in *drawing*). But the real problem with Cruse's four categories of antonymous triplet is that they are highly restrictive. Beyond the examples given, it is difficult to envisage many more antonyms which would easily fit the categories outlined without moving beyond the boundaries of what most native speakers would recognise as 'opposites'.

Evidence of this is provided by illustrations supplied by Cruse himself. For instance, as an example of a 'counteractive' triplet, Cruse cites *punch:parry:take* (1986: 202). This is a weak example because *parry* rarely relates to *punch* in text. Corpus data shows that the thing parried most commonly is a football; the 'parrier' being a goalkeeper. Furthermore, the 'opposite' of *take* would usually be identified as *give*, a fact obscured by this triplet. It would seem that Cruse has struggled to find valid examples of this category, perhaps because the number of triplets which reflect an aggressive action, a measure to neutralise it, and a consequence of an inability to neutralise it, is very limited. In defence of Cruse, the notion of the antonymous triplet is vital because it draws attention to the various prerequisites which need to be met by all verbal antonymous pairs.

Gradable antonymy

Gradable antonymy differs from non-gradable opposition in that one antonym is not automatically debarred by the other's application. In other words, it is possible to be neither *poor* nor *rich* in a way that it is not possible to be neither *baptised* nor *unbaptised*. Thus, *poor*/*rich* are gradable antonyms, as are the majority of everyday 'opposites' (*old*/*new*, *cold*/*hot*, *wet*/*dry*, etc.).

Because gradable antonyms are not mutually exclusive, they are readily modified (*quite happy*, *extremely happy*, *fairly happy*, etc.) and can take both comparative (*happier*) and superlative (*happiest*) form. Lyons (1977: 274) reports that this flexibility caused headaches for Plato, who could not comprehend how both tallness and shortness could be seemingly ascribed to the same object, as in the statement 'X is taller than Y and shorter than Z'. This, of course, is not paradoxical because *taller* and *shorter* are entirely relative concepts. Indeed, Sapir (1944; cited Palmer 1976: 94) suggests that, even in their base form, words such as *tall* and *short* are implicitly graded because they can only be understood in terms of being shorter or taller than another entity. For instance, a shallow lake is deeper than a shallow grave, a small elephant is bigger than a small child, and a narrow road is wider than a narrow ribbon. Indeed, it is almost impossible to define these antonyms in isolation; it is as though the corresponding noun is exclusively responsible for controlling the semantics of these adjectives.

Leech – alone among commentators in referring to gradable antonymy as 'polarity' (1974: 100) – notes that the above examples allude to an object-related norm. In addition to object-related norms, Leech also introduces the idea of speaker-related norms ('X is ugly' may hold true for one speaker, but not another) and role-related norms ('X is a good boss' may mean that X is good at being a boss, while not ruling out the possibility that X is poor at being a husband/wife). Leech concludes that 'it is largely because of this threefold variability of the norm that words such as *good* and *bad* are thought to be vague and shifting in their meaning' (1974: 110).

In any given pair of gradable antonyms, only one term is used to describe the degree of the gradable quality. For instance, we might ask how long or how wide something is without making any suggestion that the object is long or wide at all. Palmer (1976: 94) refers to this as the unmarked term, arguing that a question such as 'how short is it?' or 'how narrow is it?' is marked in that it implicitly suggests that the given object is short/narrow. Much attention has been given to the issue of markedness in antonymy; however, corpus data raises questions about whether the unmarked term is always regarded as being truly unmarked and supplies examples of sentences which indicate that it is not.[7]

It is often noted that the unmarked antonym is generally used to form the corresponding noun. For example, lowness and highness are measured in terms of *height*; narrowness and wideness in terms of *width*. These are both neutral, unmarked terms (compare 'what is the height of that building?' [unmarked]

to 'what is the lowness of that building?' [marked]). In English, it is the larger term that tends to be unmarked. However, Palmer notes that whereas we talk of a thickness gauge, a thinness gauge is used in Japanese (1976: 96).

Cruse (1986: 208), who refers to the unmarked and marked terms as being impartial and committed, respectively, develops a more advanced system. He begins by making a distinction between 'pseudo-comparatives' and 'true comparatives'. *Long* and *short* are said to be pseudo-comparatives because we can describe something as being long, but shorter than something else; *hot* and *cold* are said to be true comparatives because we cannot describe something as being hot, but colder than something else. In other words, we can say 'this piece of string is short, but it's longer than the other one', but we cannot say 'this bowl of soup is hot, but it's colder than the other one'. According to Cruse, once we give an item the property of hotness, we cannot subsequently describe it as being colder than another item.

Following this, Cruse notes that some antonymous pairs (such as *clever/dull*) have one true comparative and one pseudo-comparative. For example, according to Cruse, it is acceptable to say 'John's a dull lad, but he's cleverer than Bill', but unacceptable to say 'Bill's a clever lad, but he's duller than John'. Based on these very fine distinctions, Cruse develops three categories of gradable antonymy: Polar Antonymy (when both antonyms are pseudo-comparatives); Overlapping Antonymy (when one antonym is a pseudo-comparative and the other a true comparative); and Equipollent Antonymy (when both antonyms are true comparatives).

Inevitably, these sub-categories of antonymy are highly dependent on intuitive criteria. Consequently, it is not difficult to find examples which are more suspect than Cruse acknowledges. For example, *rude* and *polite* are classed as an overlapping pair because 'John's a rude lad, but he's more polite than Bill' is reckoned to be an acceptable statement. However, *happy* and *sad* are classed as an equipollent pair because 'I'm sad, but I'm happier than yesterday' is reckoned to be an unacceptable statement. My own intuitions are not compatible with this analysis. This does not make the analysis 'wrong', but it does illustrate the subjective nature of the criteria applied to develop these sub-classes of gradable antonymy. Regardless of how detailed and perceptive the categorisation is, logically based classes of antonymy are always susceptible to criticism that the individual intuition upon which they ultimately rest does not synchronise with every other individual intuition. With no recourse to real data, who is able to distinguish reliably between acceptable and unacceptable usage?

Reciprocal antonymy

The distinction between gradable and non-gradable antonyms is the most fundamental made by semanticists, but other word pairs exist which would be intuitively recognised as antonyms without easily fitting either category. For example, some form of opposition exists between *landlord* and *tenant*, yet they

are neither non-gradable antonyms (not being a landlord does not inherently make one a tenant) nor gradable antonyms (one cannot be more or less of a landlord/tenant than somebody else). Rather, they could be referred to as reciprocal antonyms. This is because the statement 'X is the landlord of Y' entails and is entailed by 'Y is the tenant of X'.

The majority of semanticists label this phenomenon 'converseness', although Palmer and Leech both prefer to speak of 'relational opposition'. Kempson (1977: 85) notes that if the variables X and Y are converse verbs, the statement A X B entails B Y A and the statement A Y B entails B X A. In other words, 'Jack precedes Jill' entails that 'Jill follows Jack', and 'Jack follows Jill' entails that 'Jill precedes Jack'. Other examples of reciprocal antonymy include *above*/*below*, *give*/*receive*, *borrow*/*lend* and *buy*/*sell*.

A fertile area for reciprocal antonymy is the field of kinship relations. If X is the *grandparent* of Y, then Y must be the *grandchild* of X. By the same token, *parent* is the reciprocal antonym of *child*, even though *adult* has already been cited as the non-gradable antonym of *child*. Of course, there is nothing to prohibit 'polyantonymy' – the instance of a word having more than one 'opposite' – but it is interesting to note the differences between the two senses of the word. Cruse (1986: 233) uses the following exchange to illustrate this:

Q. Are there any children of the marriage?
A. No, they are all grown up.

The answer to this question sounds odd because whereas the question refers to *children* in the sense of its reciprocal antonymy with *parents*, the 'zeugmatic' answer refers to *children* in the sense of its non-gradable antonymy with *adults*.

Another example of kinship antonymy is *husband*/*wife*. This is a reciprocal opposition because 'X is the husband of Y' entails and is entailed by 'Y is the wife of X'. Along the same lines, 'X is married to Y' entails and is entailed by 'Y is married to X'. However, *married to*/*married to* is a different kind of reciprocal antonymy from *husband*/*wife*, if only because the antonymy comprises lexical repetition. Most analysts (e.g. Leech 1974 or Palmer 1976) refer to this opposition as 'symmetrical'. Other examples of symmetrical antonymy include *beside*, *near to* and *meet*.

Cousin is a form of symmetrical opposition because if X is the cousin of Y, then Y must be the cousin of X. This is unusual in that, aside from *(grand)parent*/*(grand)child*, it is the only kinship opposition that is not gender-marked in English.[8] Most European languages also favour gender-marked kinship terms, but they tend not to be lexical antonyms, but morphologically related pairs which differ only by their final syllable. For example, whereas English has two distinct lexical items, *boy* and *girl*, Spanish has *muchacho* and *muchacha*, Portuguese has *menino* and *menina*, and Italian has *ragazzo* and *ragazza*. Similarly, *uncle* and *aunt* translate as *tio* and *tia* in both Spanish and Portuguese, and as *zio* and *zia* in Italian. Other examples from Spanish include *hermano*/*hermana* (brother/sister) and *sobrino*/*sobrina* (nephew/niece). The fact

that these pairs differ only by their inflection (usually a single letter) suggests that their antonymy is felt to be minimal. They are gender-marked but effectively remain the same word. This supports the view that antonyms are pairs of words which actually have much in common and questions whether gender-based antonyms should be regarded as antonyms at all (are men and women really 'opposites'?).

Returning to English, Palmer (1976: 99) notes that the terms *spouse* for husband/wife and the term *sibling* for brother/sister both avoid gender reference and are therefore symmetrical. However, such lexis is generally reserved for anthropological fields[9] and there are no corresponding neutral terms for *aunt*/*uncle*, *nephew*/*niece*, etc.

Peripheral cases of reciprocal antonymy include *doctor*/*patient*. If X is the doctor of Y, then Y must be the patient of X. However, the antonymy arising between *doctor* and *patient* is somehow different to that arising between, say, *over* and *under*, even accounting for the fact that one relationship is social and the other positional. One could express this by saying that an *over* requires an *under* (i.e. X cannot be over Y unless Y is under X), but a *doctor* does not necessarily require a *patient* (one could imagine a newly qualified GP as yet without a practice).[10] The difficulties associated with *doctor*/*patient* suggest that some form of further sub-classification may be necessary to distinguish those Xs which require Ys (as *borrow* requires *lend*) from those Xs which can exist without a corresponding Y (such as *learn* and *teach*).

Multiple incompatibility

Another traditional category of antonymy is that of multiple incompatibility. This is a borderline collection of antonyms, including pairs such as *summer*/*winter* and *north*/*south*. Indeed, Palmer (1976) and Jackson (1988) make no explicit mention of incompatibility in their entire discussion of antonymy. Leech (1974) and Kempson (1977) acknowledge the phenomenon, but refer to it as 'multiple taxonomy', while Carter provides perhaps the most workable definition: 'incompatibility . . . refers to relational contrasts between items in a semantic field' (1987: 19).

In some respects, multiple incompatibility is most similar to non-gradable antonymy. The non-gradable pair *female* and *male*, for example, belongs to a two-member system, such that X can never be simultaneously more than one member; *solid*, *liquid* and *gas*, by comparison, belong to a three-member system, such that X can never be simultaneously more than one member; similarly, *clubs*, *diamonds*, *hearts* and *spades* belong to a four-member system, such that X can never be simultaneously more than one member. And so on. Thus, multiple incompatibility may be seen as non-gradable antonymy extended to three or more terms. Whether such examples remain within the boundaries of antonymy is debatable.

The above examples of multiple incompatibility (the 'state of matter' system and the 'playing card suit' system) could be described as 'closed sets',

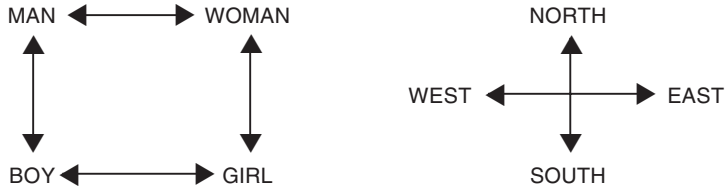
i.e. systems with a finite number of members. 'Seasons of the year' and 'cardinal compass points' are other examples. 'Open sets' of incompatibility are thinner on the ground, with the most obvious example being the numerical system. Numbers are incompatible because, for instance, if my age is twenty-seven, my age cannot simultaneously be seventeen or thirty-seven. Numbers are also an open system because there is no ceiling, no maximum digit.

Lyons (1977: 288) draws a distinction between closed sets that are 'cyclical' and closed sets that are 'serial'. An example of a cyclical set would be the days of the week, where *Wednesday* is followed by *Thursday*, *Thursday* is followed by *Friday*, and so on. The fact that *Saturday* (or *Sunday*) could be considered the end of the week is of no importance; another day will always follow. Unlike cyclical sets, serial sets always contain outermost members, an example being the letters of an alphabet. A further distinction is here made between 'scales' and 'ranks'. The illustration Lyons gives of the former is the set *excellent/good/fair/poor/bad/atrocious*. The members of this scale are incompatible in the sense that 'X wasn't just poor, he was atrocious' is a valid statement. However, I would suggest that Lyons' notion of 'scale' is too closely related to gradable antonymy to warrant differentiation. Indeed, such constructions sound odd when examples of standard multiple incompatibility are inserted: 'X wasn't just a liquid, it was a gas', etc.

The concept of 'rank' is perhaps more useful. Military rank is an appropriate example here, whereby a set ranging from *field marshal* and *general* down to *lance corporal* and *private* has been created. Again, there are similarities with gradable antonyms, but ranks appear to be more obviously incompatible with one another, and their ordered, progressive nature distinguishes them from other taxonomies. Lyons' notion of 'rank' effectively subsumes the concept of 'hierarchy', identified by Leech (1974: 106) as a fifth category of antonymy. However, Leech's sixth and final category, 'inverse opposition', warrants more attention.

For two words to be classed as inverses, the negation of one term must make the pair synonymous. This criterion is reminiscent of that for non-gradable antonymy, but Leech includes pairs such as *all/some* and *remain/become*, which may not otherwise be considered as opposites. His reasoning is that *'some* artists have no formal training' is synonymous with 'not *all* artists have formal training', and 'she will not *become* a smoker' is synonymous with 'she will *remain* a non-smoker'. This broadens the definition of antonymy in an interesting (though perhaps counter-intuitive) way.[11]

Many other word pairs have been analysed by semanticists and grouped together to create additional categories of antonymy. Though explained with care by their proponents, these categories are essentially sub-classes of multiple incompatibility and remain on the periphery of antonymy. For example, Lyons distinguishes between 'orthogonal' and 'antipodal' opposition (1977: 286). Orthogonal (meaning perpendicular, at right angles) describes the antonymy holding between the words *man*, *woman*, *girl* and *boy*. Each of these



four words contrasts with two of the other three. So *man* can be the antonym of *boy* and *woman*, but not *girl*; and *boy* can be the antonym of *girl* and *man* but not *woman*. An example of an antipodal opposition would involve the terms *north*, *east*, *south* and *west*. Here, words only contrast in one direction. So *north* is an antonym of *south*, but not *east* or *west*; and *west* is an antonym of *east* but not *north* or *south*. This is also true of the system containing the four seasons (*summer* contrasts most strongly with *winter*; *spring* with *autumn*), although it is interesting to note that Dickens contrasts 'the spring of hope' with 'the winter of despair' (p. 5).

The concept of antipodal antonymy is later discussed by Cruse (1986: 223) who proceeds to introduce many further sub-classes of antonymy including 'counterparts' (such as *hill*/*valley*, *ridge*/*groove* and *mound*/*depression*), 'independent reversives' (such as *empty*/*fill*, *enter*/*leave* and *improve*/*deteriorate*), and 'restitutives' (such as *damage*/*repair*, *kill*/*resurrect* and *stop*/*resume*). The distinction made between the latter two classes is typically subtle: *damage* is a prerequisite for *repair* (making that pair restitutives), but one does not always need to *empty* something in order to *fill* it (making that pair independent reversives).

Such categories are always interesting to consider and often provide a useful label for the pairs supplied as exemplification. However, whether one would want to identify all of these pairs as being antonymous is another question. The more restrictive the categories become, the more tenuous grows the link between the technical term 'antonymy' and the universally recognised concept of 'opposites'.

Intuitive approaches vs. data-based approaches

The section above demonstrates that 'antonymy' is a broad, inexact term traditionally used to encompass a multitude of slightly different phenomena. Semantic theorists agree that the most basic dichotomy is that between gradable and non-gradable pairs, though individual commentators rarely concur on the most appropriate terminology to describe this. However, even this most basic of divisions is challenged by Palmer, who wonders whether pairs such as *honest*/*dishonest* actually meet non-gradable as well as gradable criteria – can one really be *not honest* without being *dishonest?* (1976: 97).

This is the major problem with existing categories of antonymy: they are entirely dependent on the intuitions of their proponents. Admittedly, these intuitions are often sophisticated and usually reflect a wider consensus.

However, corpus data helps to eliminate the element of chance by tapping into not one mental lexicon, but thousands. Unlike traditional semantic theory, a corpus-based approach can also account for the fact that language is dynamic and evolving – statements unacceptable yesterday may be commonplace tomorrow. Indeed, the advantages of a textual approach are magnified when smaller, sub-categories of antonymy are examined. These sub-categories often rest on intuitive judgements about acceptability which are so subtle that uncertainty and ambiguity is unavoidable. Furthermore, the definition of antonymy applied by analysts becomes highly elastic at times in order to accommodate word pairs which, though always expressing some degree of contrast, would rarely be identified as 'good opposites' by native speakers.

In other words, to approach antonymy from an entirely intuitive position allows room for endless classification and sub-classification. However, as subsequent chapters will demonstrate, to approach antonymy from a corpus-based perspective allows room for even more classification and sub-classification. The attraction of a textual approach to antonymy is not that it results in simplification, but that it enables a set of classes to be derived which are grounded in something more solid than the shifting sands of human intuition.

Antonymy and corpora

The corpus-based research undertaken here is not the first to be conducted in the area of antonymy. Various Ph.D. theses have investigated aspects of antonymy using corpora (e.g. Murphy 1994 and Muehleisen 1997) and other researchers have made statements about antonymy based on statistical and textual evidence (e.g. Fellbaum 1995 and Kwon 1998). However, the first analysis of antonymy to make use of corpus data was published by Justeson and Katz in 1991. Three years later, Mettinger published a book about the role of antonymy in text. The work of these researchers remains the most important in establishing how antonymy operates in text and their findings will now be outlined in turn.

Using antonyms identified by Deese and (primarily) the one-million-word Brown corpus, Justeson and Katz succeed in proving that 'adjectives do indeed tend to occur in the same sentence as their antonyms far more frequently than expected by chance' (1991: 142). This is an important statement as it confirms, probably for the first time, that antonyms co-occur and that, as such, antonymy is a syntagmatic as well as a paradigmatic phenomenon. Justeson and Katz also note that antonyms 'occur in parallel and often essentially identical phrases' (1991: 142) and though no system of classification is subsequently presented, some 'typical sentences' are recorded. These contexts[12] generally involve what is termed 'substitution': the repetition of antonyms within a given framework, such as when modifying the same noun. Their examples include references to 'dry stock and wet stock' and 'a big newt and a little newt'. The conclusion reached by Justeson and Katz is that:

> The analysis of text provides not only a picture of how antonyms are used, but of what antonyms are. Our study of the appearances of antonymous adjectives in large corpora has shown that an intrinsically textual feature, the rate of phrasal substitution of antonyms for one another, is crucial to the understanding of what the relation of antonymy is, providing a lexical criterion for assessing just which word pairs are antonyms.
>
> (Justeson and Katz 1991: 184)

Therefore, Justeson and Katz advocate a description of antonymy based on textual function. This is a somewhat circular approach to the problem of definition,[13] but some reference to context is inevitable (and desirable) if a standard definition of antonymy is ever to emerge.

Mettinger (1994) adopts a more structuralist perspective than Justeson and Katz, but agrees that antonymy has always been tackled from a theoretical position and seeks to redress the balance.

> The majority of studies concerning themselves with this topic [antonymy, or 'binary semantic opposition', as Mettinger prefers] are highly theoretical in nature, thus arriving deductively at classifications and subclassifications of binary semantic opposition into various types, without, however, considering an appropriate amount of data.
>
> (Mettinger 1994: 1)

This statement is irrefutably true, but Mettinger's notion of an 'appropriate amount of data' is curious. He proceeds to examine antonymy using a corpus made up of forty-three novels, thirty of which are written by Agatha Christie. Inevitably, Mettinger's results are limited by the unusual nature of this corpus (as illustrated by the number of references to *Poirot* and *Miss Marple* in his examples!). This is naturally occurring language, but it is also modern fiction, and many of Mettinger's sentences feel stylised and literary. Another problem is corpus size – Mettinger does not specify the exact number of words in his corpus, but I would estimate a size of around three million, small in comparison to many modern corpora.

Despite its limitations, Mettinger's research remains a valuable and insightful exploration of antonymy which proves conclusively that the phenomenon is receptive to categorisation based on textual evidence. Indeed, Mettinger is able to allocate 99 of the 161 antonymous pairs examined to one of nine frames identified, simply known as frame A, frame B, etc. (1994: 40). However, Mettinger's definition of 'binary semantic opposition' is broad – included in his examples of 'Oppositions in Context' (1994: 169) are the pairs *cat/mouse*, *justice/mercy* and *murder/suicide*. While these pairs contain some element of contrast, most speakers would have reservations about labelling them as 'opposites'. Also, unlike the examples analysed in later chapters, not all of Mettinger's word pairs occur in the same sentence. This is by no means a criticism of his methodology, though it does raise questions about how close

two words must be in text to hold a relationship of significance. Justeson and Katz argue that 'the sentence is the primary rhetorical locus of [antonymous] repetitions and is the focus of our work' (1991: 140); it is the focus of my work too, but I fear that has more to do with methodological convenience than primary rhetorical loci!

Some of Mettinger's examples are similar to those identified in this investigation, but his work is not primarily classificatory. He describes his investigation as 'a first attempt at showing that contrast is a phenomenon showing greater regularity than has hitherto been assumed' (1994: 46). As such, his research is a success, though this book develops textual classifications of antonymy further and provides a more in-depth investigation of more up-to-date usage.

Other recent studies have also provided useful clues about the nature of antonymy. For example, though Murphy (1994) acknowledges that her work 'differs from other major works on antonymy in that its aim is not descriptive, nor does it originate from a structural semantics perspective' (1994: 4), the cognitive science approach which she adopts sheds light on the psycho-linguistic role of antonymy and its fellow sense relations. Also, Muehleisen (1997) investigates what makes two words antonyms and concludes that 'shared semantic range' is all-important. As an illustration, she examines the collocational profile of *wet*, *moist*, *damp*, *humid*, *dank*, *dry*, *arid* and *parched*, arguing that *dry*/*wet* are the most established antonymous pair because they have the highest frequency and share the most collocates. Finally, Fellbaum (1995) takes issue with Justeson and Katz's view that antonyms are learnt through frequent exposure to co-occurrence within particular syntactic frames, arguing that antonyms are not substitutable for one another. She makes perceptive statements about the frequency of antonymous co-occurrence, but acknowledges that her corpus is 'very small by today's standards' (1995: 286), containing only one million words. Such research contributes further to our understanding of antonymy, but this investigation benefits from a corpus much larger than the Brown corpus upon which many recent studies have been based.

Antonymy then and now

Where there is human language, there is antonymy. And where there is antonymy, there is the problem of definition. Semantic criteria alone are insufficient to explain why certain words are immediately recognised as 'opposites' while other words, equally opposed, never achieve this status. Any definition of antonymy should be lexical as well as semantic.

A second problem concerns classification. Traditional semantic theorists are notoriously poor at agreeing on terminology for categories described, but almost all commentators draw attention to a fundamental distinction between antonymous pairs: some are gradable and others are non-gradable. This distinction is apparently upheld in spite of the fact that non-gradable antonyms

are commonly graded in text. Leech (1974) and Kempson (1977) draw heavily on componential analysis to make explicit the relationship between antonyms, but the usefulness of this approach varies according to the nature of the pairs under scrutiny. More relevant is the notion of the antonymous triplet, though the problem with sub-categories developed by Cruse (and, indeed, those sub-categories developed by other semanticists) is that they are primarily based on intuitive criteria. Attempting to discern manually between acceptable and unacceptable contexts is always dicey. Moreover, it is unnecessary in the corpus age, and more recent studies of antonymy, most notably Justeson and Katz (1991) and Mettinger (1994), have made important statements about co-occurrence and demonstrated that antonymy is receptive to text-based classification.

New advances in corpus technology enable antonymy to be re-examined using real data on a much larger scale than has previously been possible. The phenomenon will now be revisited in a fresh light, with the objective of developing a new definition and new classes based on what antonymy does rather than what antonymy is. However, I agree wholeheartedly with Biber's view that 'corpus-based analysis should be seen as a complementary approach to more traditional approaches' (Biber *et al.* 1998: 9) and it should be noted that the new system of classification presented here is designed to provide an alternative to, not displace, the current body of literature about antonymy.

3 Approaching antonymy afresh
Issues of data and methodology

Existing classes of antonymy provide a useful mechanism for making logical distinctions between different word pairs, but they are not sufficient to describe the often complex ways in which antonyms operate in text. For this reason, a complementary, data-based system of classification has been developed. However, an empirical system is not inherently better than a logical system; corpora simply provide evidence upon which fresh observations may rest. If these observations are to be accurate, it is essential that the corpus is interrogated in a methodologically sound fashion. Antonymy is no more problematic to examine than other linguistic phenomena, but a number of methodological issues arise when corpus data is used. This chapter will explore these issues, justify the approaches taken, and report on each stage of the classification process.

Creating the database

When looking at 'opposites' within a corpus, one is immediately struck by the volume of contexts in which an antonymous pair co-occur. Even in a one-million-word corpus, a high-frequency pair such as *new/old* can be expected to appear in about 1,500 sentences. Therefore, in order to examine the functions of a wide range of antonymous pairs in a much larger corpus, some sort of judiciously sampled database was required. This database needed to be an accurate reflection of the way in which antonyms function across an entire corpus, and the corpus, in turn, needed to be an adequate reflection of the way in which language operates in general. Only if the corpus was chosen carefully and sampled fairly could a database be created which would allow new classes of antonymy to be safely developed.

Building a database of suitable sentences is not straightforward and questions arise at every juncture. Which corpus should be used? Which antonymous pairs should be examined? How many sentences should be included? Which sampling procedure is most appropriate? And so on. These issues will be addressed presently, but the over-riding intention was to design a database customised to meet the primary objectives of this study, namely:

- to investigate and quantify the intra-sentential functions served by antonymy in written text;

- to generate co-occurrence statistics and textual profiles for individual antonymous pairs;
- to examine variables (word class, gradability, etc.) which might affect the function of antonymy.

In order for objective 1 to be met, the database needed to be as large as possible – valid statistical conclusions could not be drawn from a small pool of sentences. However, the database could not be too large to prevent all sentences from being classified manually. With these parameters in mind, I decided that an appropriate size for the database was 3,000 sentences.

Objective 2 required this study to concern itself with a manageable number of different antonymous pairs, so that each pair could be investigated in sufficient detail. This placed restrictions on the way in which the database could be built. For example, if the first 3,000 corpus sentences to feature an antonymous pair had been sampled, it is likely that hundreds (perhaps thousands) of different word pairs would feature in the database. These pairs would be diverse in nature and even the most frequent would only arise in a few dozen examples. Therefore, it was decided to take a reverse approach and pre-specify which pairs should be studied and in which quantities: fifty-six antonymous pairs would be selected for analysis, allowing for an average of about fifty examples to be retrieved for each.

Finally, objective 3 required a number of variables which affect antonymy to be available for analysis. For example, in order to discover whether the function of antonymy differs according to word class, the database needed to include sentences which feature antonymous nouns, verbs and adverbs as well as adjectival pairs. Similarly, word pairs chosen needed to be both gradable and non-gradable, and both morphological and non-morphological.

Which corpus?

The corpus selected for use in this research needed to be large enough to yield as many contexts as possible in which pairs of antonyms (including low-frequency pairs) co-occur. With this in mind, I chose a corpus that consists of about 280 million words of text from *The Independent*. All stories printed in the newspaper between 1 October 1988 and 31 December 1996 are included in this corpus. Newspaper corpora are suitable for studies of this nature because they are large, genre-specific[1] and reflect a natural, modern, non-fictional use of written language. Thus, an overview of how antonymy is used in the field of broadsheet newspaper journalism is possible, although it should be acknowledged that antonymy might be found to function differently in other corpora.

Which pairs?

Selecting a representative sample of antonymous pairs creates more difficulties than selecting an appropriate corpus. As discussed in the previous chapter,

no single definition of antonymy has been universally agreed upon. This makes sampling problematic: how can one be sure that the antonyms selected are, indeed, genuine antonymous pairs?

In many respects, this question is impossible to resolve satisfactorily. To follow one definition of antonymy at the expense of other definitions would leave this study open to the criticism that it had not, in fact, tackled antonymy at all, but rather a specific sub-set of the phenomenon. One could also argue that such an approach would be circular if the pairs chosen pre-empted the outcome of the study. Indeed, it is difficult to imagine a list of antonyms which would not raise a single eyebrow, either because of pairs included but not considered to be 'good opposites', or because of 'good opposites' which might be conspicuous by their absence from the list.

Before explaining the strategy used here for selecting a suitable sample of antonymous pairs, the lists of antonyms which other corpus researchers have used in their investigations will be briefly considered. First, the validity of the antonymous pairs identified by Deese (1964) will be evaluated, then Mettinger's (1994) idea of using *Roget's Thesaurus* as a resource for antonyms will be discussed.

Deese's antonyms

Using the results of psycholinguistic elicitation tests as data, Deese (1964: 347–57) identified forty word pairs which he considered to be among the most fundamental in English. Justeson and Katz made use of these antonyms in their research, arguing that they are 'historically important' (1991: 142). Other studies which used the Deese antonyms include Grefenstette (1992), and Collier, Pacey and Renouf (1998), when testing automatic retrieval technology. But do the pairs listed by Deese have any claim to be the most representative of antonymy?

Being conducted before access to corpora was possible, Deese's work was based entirely on the results of word association tests. Deese took 278 adjectives[2] and used them to elicit responses from 100 informants. When a pair of contrast words successfully elicited one another more than any other word, they were added to the list of antonymous pairs, which ultimately numbered forty.

Most of these forty word pairs correspond with antonyms that native speakers would readily identify as being 'good opposites' (*bad/good*, *cold/hot*, *high/low*, etc.). However, Deese also identified other pairs which are less ingrained in the mental lexicon. For example, *alone/together* lacks the familiarity of many 'good opposites', as does *pretty/ugly*, perhaps because *beautiful* is no less an antonym of *ugly* than *pretty*. It would appear that the criterion for inclusion (that both members of an antonymous pair should elicit one another more frequently than they elicit any other term) was very broad. Though all antonyms cited by Deese fulfilled this requirement, some passed the test with alarmingly low scores. For example, given the stimulus *together*, only 6 per cent

of informants replied *alone*; given *alone*, only 10 per cent replied *together*. This was evidently enough to make these answers more popular than any other, even though the fact remains that a minimum of 84 per cent of informants failed to give either *alone* as a response to *together,* or *together* as a response to *alone*. Indeed, of the 278 adjectives tested, only one word succeeded in eliciting its antonym on a majority of occasions (*left*, to which 51 per cent of informants gave *right*).

Therefore, though there remains a strong tendency for informants to provide contrast words as responses to given stimuli in word associations tests, it may not be wise to treat Deese's forty antonyms as being in any sense exhaustive or definitive. Aside from question marks about the criteria used to develop this list (and, indeed, about how much word association tests actually reveal about the mental lexicon), Deese's 'adjectives only' policy disallowed all antonymous pairs belonging to other word classes. Furthermore, some high-frequency pairs (such as *private*/*public* and *female*/*male*) failed to qualify for Deese's list, yet many low-frequency pairs (such as *sweet*/*sour* and *married*/*single*) were included. Of course, Deese himself had no influence over which words did and did not qualify – he even acknowledged that the list 'comes nowhere near exhausting the possible independent contrasts in the language' (1964: 355) – but any study which makes use of this set of antonyms should be aware of its limitations. Deese's investigation may well be 'the standard psychological work on antonymy' (Justeson and Katz 1991: 142), but the list of antonyms that it generated is not necessarily ideal for contemporary research.

Roget's antonyms

Created in the middle of the nineteenth century, *Roget's Thesaurus* attempted to catalogue language, not in alphabetical order, but according to 'ideas'. This is of relevance to a study of antonymy because Roget chose, where possible, to present these ideas in opposition to one another. Thus, the thesaurus begins by listing words associated with *existence*, then considers words associated with *inexistence*. Following are *substantiality* and *insubstantiality*, then *intrinsicality* and *extrinsicality*. As Roget explained:

> For the purpose of exhibiting with greater distinctness the relations between words expressing opposite and correlative ideas, I have, whenever the subject admitted of such an arrangement, placed them in two parallel columns on the same page, so that each group of expressions may be readily contrasted with those which occupy the adjacent column, and constitute their antithesis.
>
> (original introduction to *Roget's Thesaurus* 1952: 545)

Mettinger made use of Roget's language organisation system as a primary source of antonymous pairs, though he acknowledged that they 'contain a number of

lexical items that are hardly used in contemporary English' (1994: 94). These would include pairs such as *fetor/fragrance*, *approbation/disapprobation* and *insalubrity/salubrity*, none of which, unsurprisingly, yielded any contextualisation in Mettinger's corpus of modern fiction. However, in addition to providing a number of established antonymous pairs (*failure/success*, *right/wrong*, *man/woman*), *Roget's Thesaurus* also supplied some interesting, 'non-core' pairs such as *courage/cowardice*, *assent/dissent* and *modesty/vanity*.

In addition to thesaural antonyms, Mettinger also looked at 'opposites in context' (1994: 169): contrast words for which he had an 'intuitive feeling that they ought to be regarded as 'opposites' in one way or the other' (1994: 2). These include very context-specific items such as *answer/ask*, *listen/look* and *gay/sad* which would not be considered 'good opposites' by many speakers, but which could reflect opposition in a given context.[3] This broadened the scope of Mettinger's study further.

However, neither using the Deese antonyms nor turning to thesaural listings is ideal. Essentially, one is still dependent on the intuitions of others to identify antonymous pairs. In the case of Roget, these intuitions are 150 years out of date; in the case of the Deese antonyms, one is reliant on the criteria for antonymy established by 1960s schools of psychology. However, it is impossible to rely on anything other than intuition when it comes to a psycholinguistic phenomenon such as antonymy. No exhaustive list of antonyms will ever be produced because the process which gives a pair of words antonymous status is complex and dynamic. Indeed, this status can only really be gauged by consensus, as definitions of antonymy vary not only from one linguist to the next, but also from one mental lexicon to the next. Hence, any list of antonyms is immediately and inherently flawed; the best one can do is to investigate a wide range of pairs which a majority of speakers might recognise as being 'good opposites'.

New index of antonyms

Given the problems associated with using established lists of antonymous pairs, I chose to create a new index of antonyms based largely on my own intuition. Some overlaps arose between this list and those antonyms identified by both Deese and Roget (it would be worrying, indeed, if this were not the case), but in creating a fresh index, I was able to discard pairs which were felt to be lower down the scale of antonymity, and include pairs which have recently achieved greater antonymous status (*gay/straight*, for instance) or have been overlooked in earlier studies for one reason or another. In total, fifty-six different antonymous pairs were selected, enough to allow for the database to be adequately representative without stretching the definition of antonymy too thinly. The over-riding criterion behind the selection of these pairs was personal intuition (these are antonymous pairs which I felt would be widely accepted as 'good opposites') but other factors were also taken into consideration:

- Despite its weaknesses, the list of antonyms created by Deese was useful as a first step towards compiling a new index because it included a number of pairs which any study of antonymy would be remiss to ignore. Of the forty adjectival pairs identified by Deese, sixteen are retained. These are mostly core items such as *poor*/*rich*, *right*/*wrong* and *hard*/*soft*.
- Only one of the antonymous pairs identified by Deese is non-gradable (*alive*/*dead*), so other non-gradable pairs were added to the new index, namely *female*/*male*, *illegal*/*legal*, *correct*/*incorrect*, *false*/*true* and *married*/*unmarried*.
- Deese's list of antonyms is restricted to adjectives, so antonymous verbs[4] (such as *confirm*/*deny*, *lose*/*win* and *hate*/*love*), adverbs (such as *explicitly*/*implicitly*, *rightly*/*wrongly* and *badly*/*well*) and nouns (such as *fact*/*fiction*, *guilt*/*innocence* and *peace*/*war*) were included in the new index.[5]
- Finally, Deese's list was restricted to lexical antonyms, so a number of morphologically related pairs (such as *advantage*/*disadvantage*, *correct*/*incorrect* and *dishonest*/*honest*) were added.

One criterion which was not used is that of frequency. Though it is true that most high-frequency antonymous pairs were included in the sample, a number of low-frequency pairs also featured in the new index. In other words, although *new*/*old* (which co-occur in 9,426 corpus sentences), *private*/*public* (6,741) and *bad*/*good* (4,804) were included, so too were pairs such as *dishonest*/*honest* (which co-occur in just 28 corpus sentences), *explicitly*/*implicitly* (32) and *officially*/*unofficially* (33). To have focused exclusively on, say, the fifty-six highest-frequency antonymous pairs in English would have resulted in a database comprised almost exclusively of 'vanilla' antonyms – gradable, non-morphological, adjectival pairs. According to the criteria outlined above, an index of fifty-six antonymous pairs was created (see Table 3.1).

Any native speakers could be reasonably expected to identify the antonym of all 112 words recorded in Table 3.1. Some ambiguity may arise because *old* contrasts with both *new* and *young*; and *right* could similarly elicit *left* as well as *wrong*. It is also possible that some words might be paired with alternative antonyms. For example, *end* could elicit *start* (as well as *begin*), *easy* could elicit *hard* (as well as *difficult*), and *happy* could elicit *unhappy* (as well as *sad*). However, this is simply a reflection of synonymy, homonymy and polysemy among antonyms; it does not suggest that the pairs selected are not all valid oppositions.

Which sampling method?

Having established which word pairs were to be analysed, the next decision concerned the number of sentences that should be sampled for each. Ideally, an equal number of sentences for every antonymous pair would have been included in the database. However, not all of the antonymous pairs listed in Table 3.1 co-occur in enough corpus sentences for an equal sample to be

Table 3.1 Antonymous pairs in the database

active/passive	advantage/disadvantage
agree/disagree	alive/dead
attack/defend	bad/good
badly/well	begin/end
boom/recession	cold/hot
confirm/deny	correct/incorrect
difficult/easy	directly/indirectly
discourage/encourage	dishonest/honest
disprove/prove	drunk/sober
dry/wet	explicitly/implicitly
fact/fiction	fail/succeed
failure/success	false/true
fast/slow	female/male
feminine/masculine	gay/straight
guilt/innocence	happy/sad
hard/soft	hate/love
heavy/light	high/low
illegal/legal	large/small
long/short	lose/win
major/minor	married/unmarried
new/old	officially/unofficially
old/young	optimism/pessimism
optimistic/pessimistic	peace/war
permanent/temporary	poor/rich
private/public	privately/publicly
punishment/reward	quickly/slowly
right/wrong	rightly/wrongly
rural/urban	strength/weakness

mathematically possible. Besides, this sampling method would have allowed for no correlation at all between the frequency of each pair and their visibility in the database. Antonyms which co-occurred in 3,000 corpus sentences would have had no greater representation than antonyms which co-occurred in 30 sentences.

A second option was to use a proportional sample. In total, 55,411 corpus sentences were found to feature both members of one of the fifty-six antonymous pairs under scrutiny. As only 3,000 were needed for this sample, 5.4 per cent of sentences in which each pair co-occur could have been retrieved. Unfortunately, this strategy would have resulted in over 500 sentences being retrieved for the highest-frequency antonymous pair, but only one or two sentences being retrieved for lower-frequency items. A dozen pairs, when sampled proportionately, would have yielded five sentences or fewer. Little of interest could be gleaned about an antonymous pair from such a tiny sample.

Ultimately, it was decided that the only way to be sure of keeping the

database balanced was to select the number of sentences to be sampled for each pair personally. Given that the pairs themselves were chosen largely according to intuitive criteria, there is no obvious reason why the number of sentences retrieved for each pair should not also be determined manually. However, some self-imposed guidelines were followed, namely that:

- no more than 60 per cent of database sentences should feature adjectival antonyms; at least 10 per cent should feature antonymous nouns, at least 10 per cent should feature antonymous verbs, and at least 10 per cent should feature antonymous adverbs;[6]
- at least 250 database sentences should feature non-gradable antonyms;
- at least 250 database sentences should feature morphologically related antonyms;
- where possible, while still meeting the criteria above, sample size should reflect co-occurrence frequency.

Using these guidelines, the number of sentences to be retrieved for each pair was determined. The representation of antonymous pairs in the database is listed in Table 3.2, in order of sample size. This table provides a breakdown of how each of the fifty-six pairs contribute to a total of 2,844 database sentences. The remaining 156 sentences will be accounted for presently, but it should first be noted that this sample has allowed for each of the specified database guidelines to be met. The majority of sentences retrieved feature antonymous adjectives, but enough feature nouns, verbs and adverbs to justify comparison. Similarly, the quotas for non-gradable pairs and morphological antonyms have been surpassed. Furthermore, a general correlation still arises between sample size and corpus frequency. For example, the highest co-occurrence rate recorded by any pair is that of *new*/*old*, and *new*/*old* features most prominently in the database with a total of 254 sentences. The second-highest co-occurrence rate recorded by any pair is that of *private*/*public*, and *private*/*public* is runner-up in the database with 134 sentences.

When a pair is more strongly represented in the database list than raw co-occurrence frequency would allow, it is usually because that pair successfully meets one or more of the other criteria. For example, *directly*/*indirectly* co-occur in a total of 492 sentences, 79 of which are included in the database. Compared with, say, *old*/*young*, which co-occur in 2,704 sentences but are only sampled 69 times, the sample size of *indirectly*/*directly* may seem proportionally large. The explanation for this is twofold: first, *indirectly*/*directly* is a morphological pair; and second, *indirectly*/*directly* is an adverbial pair. Both of these factors contribute to 'up-weight' the database representation of *directly*/*indirectly*. Conversely, some pairs were 'down-weighted' – their sample size is small relative to that of their corpus frequency because they are adjectival, gradable and lexical. Such pairs include *high*/*low*, *hard*/*soft* and *dry*/*wet*.

Having determined which antonyms were to be included in the database (and in which proportions), the procedure used to collect the 2,844 sentences

Table 3.2 Representation of antonymous pairs in the database

new/old	254	fact/fiction	36
private/public	134	strength/weakness	35
bad/good	117	confirm/deny	34
hate/love	104	gay/straight	33
poor/rich	102	hard/soft	32
active/passive	96	high/low	32
failure/success	88	illegal/legal	31
female/male	87	married/unmarried	31
directly/indirectly	79	dry/wet	31
heavy/light	77	explicitly/implicitly	30
old/young	69	attack/defend	30
feminine/masculine	68	discourage/encourage	28
fail/succeed	63	fast/slow	28
false/true	62	quickly/slowly	28
right/wrong	60	permanent/temporary	28
cold/hot	59	difficult/easy	27
lose/win	58	major/minor	27
alive/dead	54	officially/unofficially	25
badly/well	53	rural/urban	24
begin/end	51	boom/recession	24
large/small	50	optimism/pessimism	21
agree/disagree	49	punishment/reward	19
optimistic/pessimistic	47	correct/incorrect	18
privately/publicly	47	drunk/sober	18
happy/sad	45	peace/war	15
guilt/innocence	44	disprove/prove	14
rightly/wrongly	44	dishonest/honest	12
advantage/disadvantage	36		
long/short	36	**Total**	**2,844**

was straightforward. Every *n*th sentence was sampled, where *n* equals the total number of sentences in the corpus featuring both antonyms divided by the number of examples to be included in the database. For example, 2,556 corpus sentences featured both *female* and *male*, of which 87 were required. Therefore every thirtieth sentence was extracted[7] from the corpus and included in the database because 2,556 divided by 87 is approximately 30. Occasionally, a sentence was retrieved which featured both members of an antonymous pair, without those words actually operating in an antonymous relationship. Three examples from the corpus are recorded below:

2a The **new** regulations will include a ban on keeping calves in individual pens after they are eight weeks **old**, and bring in larger crates for calves up until that age.

2b The US economy is growing strongly – which means the US will

consume more oil – and by the back **end** of this year the greater demand from recovering European economies will **begin** to show.

2c Rangers, responsible for one of only two Newcastle home defeats, strove **hard** and successfully after falling behind in the fourth minute to a **soft** goal, but where it counted most the visitors had that extra touch of class.

The antonymous pair of sentence 2a is *new* and *old*, but these words clearly hold no semantic relation in this context. The fact that *new* and *old* sometimes function antonymously is irrelevant; here, their co-occurrence is simply coincidental. Likewise, sentence 2b features the antonyms *begin* and *end* in an entirely non-antonymous relationship. However, sentence 2c is slightly less clear-cut. At first glance, it seems unlikely that the adverbial use of *hard* is in any way related to the adjectival use of *soft*. But it is not impossible that the author of this sentence is matching *strove hard* with *soft goal* at some level, conscious or otherwise. This illustrates the problems of sifting valid antonymous sentences from sentences which happen to feature two words which, in other contexts, might be contrastive. When question marks arose about the validity of sentences, or when antonyms were clearly in no contextual relationship, the sentence was rejected and replaced by the next in the corpus.

Finally, an explanation of why Table 3.2 records a sample size of only 2,844: the reason for this shortfall was to allow the database to be supplemented with 156 *un*-word sentences.[8] These are sentences which feature a lexical item and the same lexical item prefixed with *un*. As *un* is the most prolific morphological marker of opposition in English, this strategy was highly productive. Three of the 156 *un*-word sentences are recorded below:

3a They are wives in limbo, because the **unknown** is always worse than the **known**.

3b If you read Sons and Lovers (**expurgated** or **unexpurgated**) you will find the midwife and her family treated with great disdain.

3c But, above all, he revered his pupils, all of them, the **clever** and the **un-clever**.

The antonymous pairs which feature in the above sentences would be on few people's lists of 'good opposites', but they make for very interesting analysis. Sentence 3a features the antonyms *unknown* and *known*, which are relatively familiar, but sentence 3b features the less everyday pair, *expurgated/unexpurgated*. Examining pairs such as the latter is useful because, although these words are quite rare, one only needs to know what *expurgated* means in order to glean the meaning of *unexpurgated*. Thus, as with all morphological antonyms, these words are immediately recognisable as 'opposites' because of their etymology, regardless of their exposure in language.[9] Similarly, the morphological antonym of sentence 3c is *un-clever* (hyphenated, one assumes,

to signal that it is a nonce word), which may not warrant a dictionary entry, but, in many respects, does not require one, as its meaning is so easily inferred.

These 156 sentences completed the database, which ultimately comprised 3,000 examples of intra-sentential antonym co-occurrence. Any sample of antonymy can be criticised on the grounds of it not being adequately representative of the phenomenon. Therefore, all subsequent claims about antonymy are based only on the following: an equitably sampled database featuring 3,000 sentences and fifty-six word pairs, all of which would be intuitively recognised as 'good opposites', to a greater or lesser degree.

Classifying the database

To describe the procedure by which all 3,000 sentences were individually categorised would be tiresome; instead, I propose to provide an insight into this process by recording a random selection of twenty-five database sentences and demonstrating the criteria used to establish the textual function of antonymous pairs therein.

4a Yet the real lesson of this book is that, while a person needs a particularly focused ambition to rise to anything significant in a business so relentlessly pacey as the media, those who **succeed** more than they **fail** are all different in the way they harness their ambition, that they have very little in common.

4b 'We are striving for the withdrawal to facilitate the re-establishment of **peace**, not **war**', he added.

4c But the gap between **rich** and **poor** has widened and there's a dwindling middle class.

4d The skills for looking after young children should be used to develop ways to help children understand the difference between **acceptable** and **unacceptable** behaviour.

4e However, Montupet's chairman, Stephane Magnan, stressed yesterday that the company's policy is to recruit **skilled** and **unskilled** workers.

4f The meticulous lawnsman will aerate it every fortnight throughout the year except when the soil is too **wet** or too **dry**.

4g Whether you **love** or **hate** ballet, Matthew Bourne's extraordinarily imaginative reworking of this classic is a thrilling feast for the senses.

4h Because of the big prices paid on world markets, much more is recovered **unofficially** than **officially**.

4i How easy to slip from the **legal** to the **illegal** trade, especially when the law is so patchy and the temptation so great.

4j *Stamps* are **popular**, but *collecting* is **unpopular**.

4k Like many of the Labour men of his day, Dan was **light** on *ideology* and **heavy** on *parish-pump politics*.

4l There was a legal principle of law that it was not permissible to blow **hot** and **cold** in the attitude one adopted.
4m 'And there's a difference', adds Forrest, 'between **good** old-fashioned and **bad** old-fashioned.'
4n Recently, I read an article by Amos Oz in the New Yorker where a majestic **old**/**new** language thundered in modern terms.
4o While his wife was **alive** he *kept silent*, but now she's **dead** he *must do what is right*.
4p All UN organisations face the same quandary – *change* **quickly** or *die* **slowly**.
4q But, **old** dog that he is, Clough has little interest in **new** tricks.
4r A study of institutionalised mental patients at the turn of the century found that **castrated** men lived longer than **uncastrated** ones.
4s The younger generation is leading the way from **passive** to **active** forms of entertainment and information-gathering.
4t He had a particular gift for metaphor, **mixed** and **unmixed**, which he deployed to great advantage.
4u However, Mr Heseltine's voter appeal towers over that of any other contender, **declared** or **undeclared**, for the leadership.
4v If you look at **employment**, not **unemployment**, that too fell in the first quarter of the year.
4w The survey also shows that the environmental movement has also won the debate over **public** versus **private** transport.
4x He hit the company right at the nadir of its fortunes, as it was trying to make the difficult and costly transition from **old** to **new** technology.
4y A private individual, not vastly **rich** and not very **poor**, dare not challenge his insurance company in the county court – supposedly inexpensive – because if he won there, the company could then drag him to the Court of Appeal, and possibly to the House of Lords, at potentially disastrous effect.

The sentences above are typical of those included in the database. For example, sentence 4a is one of the sixty-three database sentences featuring *fail* and *succeed* that were randomly retrieved from the 131 corpus sentences in which those two words co-occur. Here, *succeed* and *fail* belong to the same rank-shifted noun phrase: *those who succeed more than they fail*. Within this noun phrase, the antonyms are being measured against one another; we are being told about a group of people who both *succeed* and *fail*, but do so in uneven proportions. Thus, some form of comparison is taking place. One could say that the function of antonymy in this sentence is to pinpoint a specific class of people by comparing two success-measuring antonyms.

A similar function is served by the antonymous pair of sentences 4h and 4r. The former argues that *more is recovered unofficially than officially*; the latter states that *castrated men live longer than uncastrated ones*. Both of these examples demonstrate that antonyms can be used to set up some form of comparison.

Therefore, these three sentences were attributed to the same class; even though the antonyms are verbal in sentence 4a, adverbial in sentence 4h and adjectival in sentence 4r, their textual function is alike. The suggested name for this class is Comparative Antonymy.

Sentence 4b talks about the re-establishment of *peace, not war*. A brief glance at neighbouring sentences reveals that a parallel can be drawn between this sentence and sentence 4v, which invites us to look at *employment, not unemployment*. This pair of sentences belong together because they each negate one member of their antonymous pair and use it to post-modify the other. This is essentially a rhetorical device which places greater emphasis on the first-mentioned antonym and perhaps draws attention to the rejected alternative. The suggested name for this class is Negated Antonymy.

Sentences 4c and 4d resemble one another in terms of the way in which their respective antonymous pairs function, the former being about *the gap between rich and poor* and the latter mentioning *the difference between acceptable and unacceptable behaviour*. Clearly, these two phrases are very similar, differing only in terms of their respective noun heads, which may be different lexical items (*gap* and *difference*) but which remain semantically akin. Therefore, a third textual function of antonymy can be identified. The only other context in which this function is served is sentence 4m which speaks of *a difference . . . between good old-fashioned and bad old-fashioned.* The suggested name for this class is Distinguished Antonymy.

Sentence 4e states that a company's policy is to recruit *skilled and unskilled workers*. Here, the two antonyms are presented in a unified, coordinated fashion to express inclusiveness. It matters not whether the workers are skilled or unskilled, the company's policy remains the same. Similarly, in sentence 4g, it matters not *whether you love or hate ballet*, the reworking of the classic is still a thrilling feast. Again, the antonyms are coordinated, though in this context they express exhaustiveness rather than inclusiveness. A third illustration of this textual function of antonymy can be found in sentence 4t, which notes that metaphors may be *mixed and unmixed*; once again, it does not matter which antonym is applicable here, so the pair are brought together and effectively stripped of their antonymous power. A final examples of this phenomenon can be found in sentence 4u, which refers to *any other contender, declared or undeclared*. In each of these examples, a pair of antonyms, linked by *and* or *or*, modify a noun phrase to signal inclusiveness or exhaustiveness. The suggested name for this class is Coordinated Antonymy.

At first glance, sentence 4f also seems to belong to the class of Coordinated Antonymy because antonyms are linked by *or* and no overt contrast is created. However, this reference (to soil being *too wet or too dry*) is, in fact, fundamentally different from the examples above. This is because sentence 4f modifies each of its antonyms with *too*, thereby preventing the antonymous pair from being truly exhaustive. Soil *too wet or too dry* does not equate with 'all kinds of soil' in the way that metaphor *mixed and unmixed* equates with 'all kinds of metaphor'. For the same reason, sentence 4y (which refers to

individuals *not vastly rich and not very poor*) belongs not to the class of Coordinated Antonymy, but to a new class for sentences which use extremity-signalling adverbs to incorporate either end of a semantic scale only. The suggested name for this class is Extreme Antonymy.

Sentence 4i notes how easy it is to slip *from the legal to the illegal trade*. Analogies can be drawn between this example and sentences 4s (*the younger generation is leading the way from passive to active forms of entertainment*) and 4x (*the costly transition from old to new technology*). All three contexts refer to a movement or change from one antonymous state to another. The suggested name for this class is Transitional Antonymy.

Two contrasts are presented in sentences 4j: that between the antonymous pair *popular* and *unpopular* and that between *stamps* and *collecting*. Forms of the verb *to be* link *stamps* to *popular* and link *collecting* to *unpopular*. In this way, some of the antonymity of *popular*/*unpopular* could be said to 'rub off' on *stamps*/*collecting*. This is useful because *stamps* and *collecting* would not usually be regarded as contrastive. In other words, a familiar antonymous pair is effectively acting as a lexical signal that we should interpret a non-antonymous pair contrastively. Another context that features two related oppositions is sentence 4k, in which somebody is described as *light on ideology and heavy on parish-pump politics*. In this context, *light* and *heavy* encourage us to contrast *ideology* and *parish-pump politics*. Similarly, *alive* and *dead* create a contrast between *kept silent* and *must do what is right* in sentence 4o; and *quickly* and *slowly* create a contrast between *change* and *die* in sentence 4p. The textual function of the antonymous pair in each of these examples is to instruct us to interpret a nearby pair of words or phrases contrastively. The suggested name for this class is Ancillary Antonymy.

Sentence 4l is unusual because it features the expression *to blow hot and cold*. One could make a case for this sentence to be attributed to the class of Coordinated Antonymy, but this would overlook the important fact that *to blow hot and cold* is an idiomatic expression that is unlikely to be processed in the same way as non-idiomatic language. Like most words, antonyms sometimes function as part of an idiom, and in sentence 4l (and in sentence 4q, which features a corruption of the old-dog-new-tricks proverb), the textual function of the antonymous pair is overshadowed by the idiomaticity of the phrase. Hence, a new class of antonymy is required. The suggested name for this class is Idiomatic Antonymy.

So far, twenty-three of the twenty-five sentences above have been analysed according to the textual function of their antonymous pair and classified accordingly. However, some sentences inevitably resist generalisation and two contexts need to be attributed to the class of 'others'. These are sentence 4n, in which *old* and *new* are joined by an oblique stroke, and sentence 4w, which refers to the debate over *public versus private transport*. The textual function of the antonymous pair in each of these examples is too rare to justify classification, though these sentences illustrate that antonymy is receptive to innovative and unusual usage in text.

Table 3.3 Functions of antonymy in database sample

Suggested name	*Sentences*
Ancillary Antonymy (An) *(stamps are popular, but collecting is unpopular)*	4j, 4k, 4o, 4p
Coordinated Antonymy (Co) *(policy is to recruit skilled and unskilled workers)*	4e, 4g, 4t, 4u
Comparative Antonymy (Cm) *(those who succeed more than they fail)*	4a, 4h, 4r
Distinguished Antonymy (Ds) *(the gap between rich and poor has widened)*	4c, 4d, 4m
Transitional Antonymy (Tr) *(how easy to slip from the legal to the illegal trade)*	4i, 4s, 4x
Negated Antonymy (Ng) *(to facilitate the re-establishment of peace, not war)*	4b, 4v
Extreme Antonymy (Ex) *(except when the soil is too wet or too dry)*	4f, 4y
Idiomatic Antonymy (Id) *(it was not permissible to blow hot and cold)*	4l, 4q
Others	4n, 4w

The twenty-five database sentences 4a to 4y show that the function of antonymy does not differ with every context. Some patterns of usage emerge from the data and it would seem that some of these patterns are more common than others. In total, eight different functions of antonymy have been identified, excluding the two examples which show antonyms functioning in anomalous ways. These eight classes are summarised in Table 3.3.

This exercise is intended to provide an insight into the way in which the database was classified. Only twenty-five examples have been discussed, but the remaining 2,975 database sentences have been subject to similar scrutiny and also attributed to one of the eight new classes listed in Table 3.3, where possible. So how do all 3,000 sentences distribute among these classes?

Database distribution

Table 3.4 provides a breakdown of the entire database according to the classificatory system outlined above. All of the word pairs selected for study (including the *un*-words portion of the sample) are recorded in the left-hand column. Subsequent columns record the number of database sentences in which they serve each textual function.

This table demonstrates that the way in which antonyms function in text is highly predictable. Two classes dominate: Ancillary Antonymy, to which 38.7 per cent of sentences have been attributed, and Coordinated Antonymy, to which 38.4 per cent have been attributed. These two classes are significantly

Table 3.4 Distribution of pairs across new classes of antonymy

	An	*Co*	*Cm*	*Ds*	*Tr*	*Ng*	*Ex*	*Id*	*Oth*	*Total*
active/passive	53	14	9	6	6	6	0	0	2	96
advantage/disadvantage	15	14	2	0	4	1	0	0	0	36
agree/disagree	26	17	3	0	0	0	0	0	3	49
alive/dead	16	26	9	1	0	1	0	0	1	54
attack/defend	10	15	3	0	0	2	0	0	0	30
bad/good	55	47	4	4	3	1	0	2	1	117
badly/well	31	15	4	1	0	2	0	0	0	53
begin/end	24	23	3	0	0	1	0	0	0	51
boom/recession	12	3	4	0	5	0	0	0	0	24
cold/hot	21	23	0	0	2	0	2	11	0	59
confirm/deny	0	34	0	0	0	0	0	0	0	34
correct/incorrect	6	11	0	1	0	0	0	0	0	18
difficult/easy	19	5	0	0	0	0	1	0	2	27
directly/indirectly	21	57	1	0	0	0	0	0	0	79
discourage/encourage	16	8	2	0	0	2	0	0	0	28
dishonest/honest	8	4	0	0	0	0	0	0	0	12
disprove/prove	0	14	0	0	0	0	0	0	0	14
drunk/sober	8	7	0	0	1	1	0	0	1	18
dry/wet	11	9	3	1	3	0	4	0	0	31
explicitly/implicitly	6	19	2	0	0	3	0	0	0	30
fact/fiction	5	5	2	11	2	4	1	0	6	36
fail/succeed	30	27	5	0	0	1	0	0	0	63
failure/success	38	20	10	12	1	6	1	0	0	88
false/true	10	34	3	11	0	1	1	0	2	62
fast/slow	17	7	2	1	0	0	0	0	1	28
female/male	23	43	1	4	1	0	0	2	13	87
feminine/masculine	37	18	2	3	0	0	1	0	7	68
gay/straight	3	20	7	1	1	1	0	0	0	33
guilt/innocence	5	27	3	5	1	2	0	0	1	44
happy/sad	22	17	2	0	2	0	2	0	0	45
hard/soft	17	3	2	1	3	2	3	0	1	32
hate/love	40	44	7	2	1	2	2	0	6	104
heavy/light	46	19	5	1	4	0	2	0	0	77
high/low	20	3	2	3	1	1	1	1	0	32
illegal/legal	10	17	1	0	3	0	0	0	0	31
large/small	17	23	4	2	2	0	2	0	0	50
long/short	22	7	4	1	0	1	0	1	0	36
lose/win	27	25	5	0	0	1	0	0	0	58
major/minor	11	9	0	3	4	0	0	0	0	27
married/unmarried	4	14	8	5	0	0	0	0	0	31
new/old	81	76	21	19	10	3	1	6	37	254
officially/unofficially	14	10	1	0	0	0	0	0	0	25
old/young	20	34	6	5	0	1	3	0	0	69
optimism/pessimism	30	12	3	0	1	0	1	0	0	47
optimistic/pessimistic	9	1	2	0	6	1	1	0	1	21
peace/war	3	5	1	1	1	2	0	0	2	15
permanent/temporary	6	12	5	1	3	1	0	0	0	28
poor/rich	46	16	6	24	1	0	5	0	4	102
private/public	36	68	6	13	5	2	0	0	4	134

Table 3.4 (*continued*)

	An	*Co*	*Cm*	*Ds*	*Tr*	*Ng*	*Ex*	*Id*	*Oth*	*Total*
privately/publicly	20	24	2	1	0	0	0	0	0	47
punishment/reward	6	5	4	0	0	3	0	0	1	19
quickly/slowly	16	6	2	0	0	0	4	0	0	28
right/wrong	36	13	1	5	1	0	0	0	4	60
rightly/wrongly	1	43	0	0	0	0	0	0	0	44
rural/urban	7	13	0	2	1	0	1	0	0	24
strength/weakness	11	6	6	0	4	4	0	0	4	35
un-words	58	60	15	10	7	3	1	0	2	156
Total	**1,162**	**1,151**	**205**	**161**	**90**	**62**	**40**	**23**	**106**	**3,000**
Percentage:	38.7	38.4	6.8	5.4	3.0	2.1	1.3	0.8	3.5	100

larger than any others and collectively account for 77.1 per cent of all database sentences. The third-largest class is Comparative Antonymy, but this is only a fraction the size of the two dominant classes, accounting for 205 sentences, less than 7 per cent of the database total. Distinguished Antonymy accounts for a further 5.4 per cent of database sentences, but no more than 100 sentences have been attributed to any of the other new classes. Only 3.0 per cent of the database was classified in terms of Transitional Antonymy, and the classes of Negated, Extreme and Idiomatic Antonymy account for 2.1 per cent, 1.3 per cent and 0.8 per cent of sentences respectively. In other words, the distribution of database sentences across the eight new classes identified is anything but uniform, with two major classes absorbing the bulk of sentences and the frequency of minor classes tailing off very rapidly.

Perhaps most remarkable is that the majority of individual antonymous pairs follow a roughly similar distribution, being classified primarily in terms of Ancillary Antonymy and Coordinated Antonymy, but occasionally serving some minor functions as well. Individual word pairs which reflect this global pattern include *bad*/*good*, *begin*/*end* and *hate*/*love*. Therefore, Ancillary and Coordinated Antonymy are the most popular classes, not because they are each strongly favoured by a small number of pairs, but because most 'opposites' serve these two functions so regularly. Indeed, in the case of forty-four of the fifty-six pairs sampled, Ancillary and Coordinated Antonymy both account for more sentences than any other class.

Of course, some individual antonymous pairs display idiosyncratic tendencies and favour certain textual functions at the expense of others. This is demonstrated by Table 3.5, which compares the raw frequency of the distribution for individual antonymous pairs with the average for all antonymous pairs. This enables a normal distribution score to be calculated; a very high or a very low score indicates that the antonymous pair in question strongly favours or strongly disfavours that particular textual function.

Table 3.5 confirms that the majority of individual antonymous pairs distribute across new classes in proportions similar to the distribution for all pairs

Table 3.5 Breakdown of database expressed as normal distribution score

	An	*Co*	*Cm*	*Ds*	*Tr*	*Ng*	*Ex*	*Id*	*Oth*
active/passive	0.82	0.12	0.66	0.62	0.69	0.84	0.30	0.42	0.45
advantage/disadvantage	0.57	0.49	0.41	0.25	0.91	0.52	0.30	0.42	0.28
agree/disagree	0.79	0.41	0.45	0.25	0.26	0.25	0.30	0.42	0.79
alive/dead	0.30	0.66	**0.95**	0.35	0.26	0.42	0.30	0.42	0.43
attack/defend	0.38	0.70	0.70	0.25	0.26	0.86	0.30	0.42	0.28
bad/good	0.68	0.52	0.27	0.44	0.43	0.32	0.30	0.68	0.35
badly/well	0.87	0.30	0.54	0.35	0.26	0.63	0.30	0.42	0.28
begin/end	0.68	0.61	0.43	0.25	0.26	0.44	0.30	0.42	0.28
boom/recession	0.74	0.10	**0.95**	0.25	**1.00**	0.25	0.30	0.42	0.28
cold/hot	0.43	0.49	0.12	0.25	0.49	0.25	0.73	**1.00**	0.28
confirm/deny	0.01	**1.00**	0.12	0.25	0.26	0.25	0.30	0.42	0.28
correct/incorrect	0.38	0.85	0.12	0.58	0.26	0.25	0.30	0.42	0.28
difficult/easy	**0.96**	0.16	0.12	0.25	0.26	0.25	0.76	0.42	0.86
directly/indirectly	0.25	0.94	0.17	0.25	0.26	0.25	0.30	0.42	0.28
discourage/encourage	0.85	0.30	0.52	0.25	0.26	0.89	0.30	0.42	0.28
dishonest/honest	0.94	0.39	0.12	0.25	0.26	0.25	0.30	0.42	0.28
disprove/prove	0.01	**1.00**	0.12	0.25	0.26	0.25	0.30	0.42	0.28
drunk/sober	0.63	0.49	0.12	0.25	0.64	0.79	0.30	0.42	0.75
dry/wet	0.43	0.31	0.68	0.43	0.87	0.25	**1.00**	0.42	0.28
explicitly/implicitly	0.15	0.87	0.48	0.25	0.26	**0.98**	0.30	0.42	0.28
fact/fiction	0.08	0.11	0.41	**1.00**	0.64	**0.99**	0.66	0.42	**1.00**
fail/succeed	0.69	0.57	0.57	0.25	0.26	0.40	0.30	0.42	0.28
failure/success	0.60	0.21	0.78	0.93	0.33	0.87	0.45	0.42	0.28
false/true	0.10	0.77	0.36	**0.98**	0.26	0.40	0.51	0.42	0.56
fast/slow	0.89	0.25	0.52	0.45	0.26	0.25	0.30	0.42	0.59
female/male	0.24	0.69	0.16	0.52	0.33	0.25	0.30	0.76	**1.00**
feminine/masculine	0.81	0.27	0.25	0.51	0.26	0.25	0.49	0.42	**0.96**
gay/straight	0.05	0.85	**0.99**	0.42	0.46	0.55	0.30	0.42	0.28
guilt/innocence	0.06	0.85	0.49	0.87	0.41	0.70	0.30	0.42	0.47
happy/sad	0.72	0.47	0.34	0.25	0.56	0.25	0.83	0.42	0.28
hard/soft	0.79	0.08	0.45	0.43	0.85	0.84	**1.00**	0.42	0.55
hate/love	0.49	0.56	0.49	0.35	0.32	0.43	0.55	0.42	0.77
heavy/light	0.88	0.24	0.47	0.32	0.62	0.25	0.64	0.42	0.28
high/low	0.91	0.08	0.45	0.79	0.47	0.56	0.70	0.85	0.28
illegal/legal	0.36	0.77	0.26	0.25	0.87	0.25	0.30	0.42	0.28
large/small	0.40	0.63	0.57	0.48	0.53	0.25	0.79	0.42	0.28
long/short	0.90	0.17	0.76	0.40	0.26	0.52	0.30	0.81	0.28
lose/win	0.67	0.57	0.61	0.25	0.26	0.41	0.30	0.42	0.28
major/minor	0.55	0.39	0.12	0.86	**0.98**	0.25	0.30	0.42	0.28
married/unmarried	0.07	0.61	**1.00**	**0.97**	0.26	0.25	0.30	0.42	0.28
new/old	0.35	0.33	0.59	0.70	0.53	0.36	0.35	0.76	**1.00**
officially/unofficially	0.84	0.51	0.31	0.25	0.26	0.25	0.30	0.42	0.28
old/young	0.29	0.68	0.62	0.68	0.26	0.38	0.82	0.42	0.28
optimistic/pessimistic	0.92	0.26	0.46	0.25	0.40	0.25	0.58	0.42	0.28
optimism/pessimism	0.59	0.05	0.67	0.25	**1.00**	0.72	0.86	0.42	0.69
peace/war	0.15	0.39	0.48	0.65	0.71	**1.00**	0.30	0.42	**0.99**
permanent/temporary	0.16	0.57	**0.97**	0.45	0.90	0.61	0.30	0.42	0.28
poor/rich	0.64	0.13	0.43	**1.00**	0.32	0.25	0.87	0.42	0.62
private/public	0.25	0.71	0.34	0.81	0.51	0.39	0.30	0.42	0.54

Table 3.5 (*continued*)

	An	*Co*	*Cm*	*Ds*	*Tr*	*Ng*	*Ex*	*Id*	*Oth*
privately/publicly	0.59	0.71	0.32	0.36	0.26	0.25	0.30	0.42	0.28
punishment/reward	0.34	0.27	**0.99**	0.25	0.26	**1.00**	0.30	0.42	0.73
quickly/slowly	0.85	0.20	0.52	0.25	0.26	0.25	1.00	0.42	0.28
right/wrong	0.89	0.20	0.18	0.74	0.37	0.25	0.30	0.42	0.82
rightly/wrongly	0.02	**1.00**	0.12	0.25	0.26	0.25	0.30	0.42	0.28
rural/urban	0.29	0.76	0.12	0.74	0.54	0.25	0.81	0.42	0.28
strength/weakness	0.34	0.15	**0.96**	0.25	0.92	**0.99**	0.30	0.42	**0.98**
un-words	0.47	0.48	0.68	0.63	0.57	0.43	0.38	0.42	0.38

in the database. Of the 513 normal distribution scores calculated, only 32 are statistically significant (i.e. over 0.95 or under 0.05). This indicates that most antonyms function in a relatively predictable fashion in text. However, it is interesting to explore those pairs which resist the global trend for antonyms and favour certain classes disproportionately. To take Distinguished Antonymy as an illustration, Table 3.5 reveals that four antonymous pairs serve this function at a statistically significant rate (and therefore score over 0.95): *fact*/*fiction*, *false*/*true*, *married*/*unmarried* and *poor*/*rich*. The last-named pair achieve the highest score (1.00), but this is not surprising given the political focus of *The Independent* newspaper. When *poor* and *rich* co-occur in the corpus, it is often the gap between the two which is being discussed. No antonymous pairs can be said to disfavour Distinguished Antonymy at a statistically significant rate because its proportional representation across the whole database is relatively low.

The class of Coordinated Antonymy is, according to Table 3.5, strongly favoured by three pairs: *confirm*/*deny*, *disprove*/*prove* and *rightly*/*wrongly*, all of which record a normal distribution score of 1.00. Corpus data shows that antonym co-occurrence among this trio of pairs almost always takes place in coordinated context. Because these pairs favour Coordinated Antonymy, they inevitably disfavour the other dominant class, Ancillary Antonymy. The antonymous pair which records the highest score for Ancillary Antonymy is *difficult*/*easy* (0.96), which indicates that, when these words co-occur, they are often used in connection with another pair of contrastive expressions (e.g. *X is difficult, but Y is easy*). Other clusters of antonymous pairs can be identified for each new class. For example, Extreme Antonymy is favoured by adjectival pairs such as *dry*/*wet* and *hard*/*soft*, whereas Negated Antonymy tends to be favoured by nouns such as *peace*/*war* and *punishment*/*reward*.

However, most antonymous pairs do not show anything more than a slight inclination (or disinclination) towards any particular class. The typical antonymous pair functions in terms of Ancillary and Coordinated Antonymy in approximately equal proportions, these proportions being much greater than those recorded for other functions. Because these two classes are the most widespread, each is allotted an entire chapter of this book. Ancillary Antonymy is

analysed in Chapter 4 and Coordinated Antonymy is analysed in Chapter 5. A taxonomy of all other classes is presented in Chapter 6.

The extent to which the database distribution can be regarded as an accurate reflection of the way in which antonyms function in text is entirely dependent on the methodology used to sample the corpus. Establishing which antonymous pairs are best able to epitomise the phenomenon of antonymy requires subjective choices to be made. Similarly, the process of creating and classifying a suitable database does not become an exact science simply because a corpus-based approach is taken. Corpus linguistics is not intuition-free linguistics. However, all reasonable precautions have been taken to ensure that this research is based on a methodologically sound exploitation of corpus data from which meaningful statistics may be derived and upon which safe conclusions may be drawn. The database is unlikely to be a flawless portrait of antonymous usage, but it can safely be regarded as an accurate snapshot of how 'opposites' currently operate in English.

4 New classes of antonymy I
Ancillary Antonymy

According to the distribution of database sentences presented in Table 3.1, as many as two sentences in five which feature a familiar antonymous pair also contain a second, related opposition. This means that the ancillary effect is not only common among 'opposites' in text, but is a relatively widespread phenomenon across language in general. Ancillary Antonymy will now be exemplified extensively and analysed in detail, with a view to assessing the different ways in which an antonymous pair can act as a lexical signal of a nearby contrast.

What is Ancillary Antonymy?

When first exploring the ways in which antonymy operates in text, it was striking to note the number of sentences which featured two contrasts: one arising between antonyms and one arising between another pair of words or phrases. The latter contrast was often closely related to the former, and it seemed that the antonymous pair was partly responsible for generating (or at least affirming) this new opposition. This phenomenon is evidenced in each of the following sentences. Although these examples are semantically, syntactically and grammatically distinct, it can be argued that the antonymous pair of each contributes to a larger contrast; that the antonyms themselves are not the primary contrast of the sentence, but are actually responsible for signalling a more important opposition (usually instantial) between another pair of words, phrases or clauses.

5a I **love** *to cook* but I **hate** *doing the dishes* – so I'd have a dishwasher or a family of gypsies to do the washing up.

5b Robin Cook, Labour's health spokesman, demanded: 'How can it be **right** to limit the hours worked by *lorry drivers and airline pilots*, but **wrong** to limit the hours of *junior hospital doctors undertaking complex medical treatment*?'

5c Since then, of course, they've all had knighthoods, usually when they're too **old** to play *Hamlet* but too **young** to play *butlers in Hollywood movies.*

5d *Eighty-five* per cent of 'firm Tories' **agree** that 'a Labour government would wreck the economy'; only *six* per cent **disagree**.

5e At Worcester on Wednesday, Botham – apart from bowling well – was wandering around in a T-shirt with the message: '*Form* is **temporary**, *class* is **permanent**'.

5f As the Governor of Kumomoto province told me, 'This is a **rich** *country*, with **poor** *people*'.

5g It is meeting **public** *need*, not **private** *greed*.

5h If so, unemployment may rise more **quickly** *now*, but more **slowly** *later*.

5i He also suggests discipline should be tailored differently, saying **extroverts** are most motivated by **reward** while **introverts** respond more to **punishment**.

The nine database sentences recorded above, though diverse in many respects, all illustrate the phenomenon of Ancillary Antonymy. To begin with sentence 5a, the recognised antonyms here are *love*/*hate* and these words are linked to another contrast, namely that between *to cook* and *doing the dishes*. Furthermore, one could argue that this latter contrast is the more fundamental of the two. This is because the distinction between cooking and washing up is more specific, and the fact that these expressions are not familiar 'opposites' (as *love* and *hate* are) actually seems to focus our attention more on the contrastive context in which they are placed. One could put it like this: it is not contrastive antonyms which catch the eye, it is contrastive non-antonyms. For ease of reference, in sentence 5a, we could label *love* and *hate* as the 'A-pair' (because they are both antonymous and ancillary) and *to cook* and *doing the dishes* as the 'B-pair' (because they are the second contrast of the sentence).[1]

One question which arises immediately is this: how can we tell which pair expresses the more important contrast? Is the author's over-riding intent to oppose established antonyms (presented here in bold) or to oppose concepts which might not otherwise be interpreted contrastively (presented here in italics)? The answer to this question is, I suggest, more likely to be the latter.

Consider sentence 5b, which includes a rhetorical question posed by Robin Cook. The antonymous pair here is *right*/*wrong*, but the point of the question is not to contrast these words, but rather to contrast *lorry drivers and airline pilots*, on one hand, and *junior hospital doctors undertaking complex medical treatment*, on the other. The intended effect of Cook's rhetoric is to compare the relative importance of these two sets of professions. This suggests that the B-pair (*lorry drivers and airline pilots* versus *junior hospital doctors undertaking complex medical treatment*) is the more important contrast. The A-pair (*right* and *wrong*) is one of a number of factors which contribute to the overall contrastive nature of this sentence.

The primary contrast of sentence 5c concerns acting roles and, specifically, acting roles suitable for people of a given age. The profundity of playing *Hamlet* is set up in opposition against the triviality of playing *butlers in*

Hollywood movies, a B-pair contrast signalled by the A-pair *old* and *young*. Sentence 5d differs in that its primary contrast is numeric. Here, the *eighty-five* per cent of 'firm Tories' who agree with a proposition is contrasted with the *six* per cent who do not. Once again, the focus of this sentence is the B-pair (*eighty-five* and *six*) rather than the A-pair (*agree* and *disagree*).

Each of the other sentences follows a similar pattern: *temporary* and *permanent* signal a contrast between *form* and *class* in sentence 5e; *rich* and *poor* signal a contrast between *country* and *people* in sentence 5f; *public* and *private* signal a contrast between *need* and *greed* in sentence 5g; and *quickly* and *slowly* affirm a contrast between the near-antonymous, temporal notions of *now* and *later* in sentence 5h. Sentence 5i goes one step further by featuring two pairs of words which most speakers would regard as being 'opposites' (*extroverts/introverts* and *punishment/reward*). While one cannot state with confidence which pair is the B-pair, one can safely assert that both pairs contribute to a larger contrast between their respective clauses and, as such, function in some kind of symbiotic relationship.

Thus, sentences 5a–i can be described as follows: in each context, two oppositions occur, one between an established antonymous pair (known here as the A-pair) and the other between a pair of words or phrases (known here as the B-pair) which are also intended to be interpreted contrastively, but have less innate opposition than the A-pair. What happens in these sentences is that the A-pair becomes ancillary to the B-pair and instructs us to treat the B-pair more contrastively: if the B-pair has no innate element of opposition, the A-pair generates an instantial contrast; if the B-pair already has a low level of innate opposition, the A-pair activates this latent contrastive potential; and if the B-pair already has a high level of innate opposition, the A-pair affirms this contrast to the point of assigning antonymity. In this way, the B-pair is effectively nudged further up the scale of opposition by the A-pair.

The A-pairs in sentences 5a–i are all responsible for contributing towards a larger contextual contrast. However, these sentences differ from one another in a number of ways. For example, the nature of the B-pair in each sentence is clearly not uniform: a B-pair can take the shape of single words (e.g. sentence 5e: *form* and *class*) or can be multi-word expressions (e.g. sentence 5a: *to cook* and *doing the dishes*); B-pairs can also belong to a variety of word classes (e.g. nouns in sentence 5g (*need/greed*) and adverbs in sentence 5h (*now/later*)); and they can represent a variety of semantic fields (e.g. professions in sentence 5b (*lorry drivers/junior doctors*) and acting roles in sentence 5c (*Hamlet/butlers in Hollywood movies*)). Furthermore, some B-pairs have very little inherent contrastive power (e.g. sentence 5e: *form* and *class*); while others feature a strong element of innate opposition (e.g. sentence 5h: *now* and *later*). Finally, it is notable that some Ancillary Antonymy sentences employ overt markers of contrast, while others do not. Four of the nine sentences above use *but* to link clauses, and *not* and *while* are each used once. However, not all sentences rely on contrast-signalling conjunctions – *with* appears on one occasion and a couple of examples rely only on punctuation.

However, although a number of differences between examples of Ancillary Antonymy can be identified, so too can a number of similarities. For example, parallelism plays a major role in the construction of sentences 5a–i: repetition of words, phrases and clauses contribute significantly to the contrastive nature of each context; grammatical parallelism also serves an important non-lexical role in creating textual opposition. This is illustrated by sentence 5c which compares being *too old to play Hamlet* with being *too young to play butlers in Hollywood movies*. As demonstrated below, both clauses begin with *they're too* (with ellipsis of subject and verb in clause two), then feature their respective A-pair member, then feature the infinitive *to play*, then their respective B-pair member. This parallelism is a key contrast-affirming mechanism and is evident, to a greater or lesser degree, in all sentences belonging to Ancillary Antonymy.

Sentence 5c

		A-pair		**B-pair**
clause 1:	they're too	old	to play	Hamlet
clause 2:	[they're] too	young	to play	butlers in Hollywood movies

These aspects of Ancillary Antonymy sentences will now be investigated further: first, a taxonomy of B-pairs will be presented to demonstrate the diverse nature of words and phrases which receive additional contrastive power from an A-pair; second, the levels of contrast inherent in B-pairs will be examined in order to gauge exactly how ancillary an A-pair really is; third, the role of parallelism in Ancillary Antonymy sentences will be analysed to demonstrate the importance of inter-clausal repetition; fourth, conjunctions will be examined to determine how clauses are connected together; and, finally, Ancillary Antonymy sentences which syntactically detach their B-pair from their A-pair will be exemplified to illustrate how the ancillary effect can be achieved in less orthodox fashion.

A taxonomy of B-pairs

The A-pair in an Ancillary Antonymy sentence is responsible for signalling a second contrast elsewhere in the sentence. This second contrast is between B-pairs – words or phrases which are usually co-hyponyms or co-taxonyms belonging to a given (though often context-specific) set. In this section, eight different types of B-pair will be outlined and discussed. These categories are heterogeneous and incorporate semantic, conceptual and referential criteria: the first five categories describe B-pairs whose members are co-hyponyms (of superordinate terms such as *people* or *places*) and are listed from the more specific to the more general; the final three categories describe B-pairs which are metalinguistically related to one another (e.g. they comprise a synonymous or meronymous pair). The eight categories below do not exhaust the number of relationships which members of a B-pair may hold with one another; rather, they provide an insight into the irregular nature of these relationships.

Political B-pairs

6a 'The issue at the next election will be between **fair** taxation under *Labour* and **unfair** taxation under *the Conservatives*', he said.
6b Broadly speaking, the community charge was **popular** with *Conservative* voters and **unpopular** with *Labour* voters.
6c *Communism* may be **dead**, but *fascism* is most assuredly **alive**.

As one would expect from a corpus of broadsheet newspaper text, political issues are the topic of many database sentences and the examples above illustrate how an A-pair may be used to affirm a contrast between two political parties or ideologies. For example, sentence 6a makes an association between *Labour* and *fair taxation*, on one hand, and *the Conservatives* and *unfair taxation* on the other. The cultural opposition in British politics between *Labour* and the *Conservatives* is very well established (one certainly could not argue that *fair* and *unfair* 'create' this contrast in any sense), but the A-pair seem to contribute to the overall balance of the sentence, encouraging us to read the statement as a direct contrast between the two noun phrases on either side of the conjunction *and*. Sentence 6b also contrasts *Conservative* and *Labour*, but sentence 6c illustrates that an A-pair may also signal a political contrast between other parties/ideologies. Here, *communism* and *fascism* are presented in opposition.

Human B-pairs

7a *Mrs Thatcher* has been a **lucky** prime minister, *Mr Heath* was an **unlucky** one.
7b *Charles*, **unskilfully**, is playing for the popular vote; *Diana*, very **skilfully**, is doing the same.
7c *Kennedy* **dead** is more interesting than *Clinton* **alive**.

People also figure prominently in journalistic text and the B-pair in each of the sentences above is human. In other words, A-pairs are being used to reflect/create a more instantially important contrast between two individuals. In sentence 7a, these individuals are *Mrs Thatcher* and *Mr Heath*; in sentence 7b, they are *Charles* and *Diana.* It is interesting to note that these B-pairs are not only co-hyponyms of human beings, but also co-hyponyms of a more specific superordinate: *Mrs Thatcher* and *Mr Heath* are both politicians (indeed, as the text notes, they are both former prime ministers) and *Charles* and *Diana* both belong to a specific sub-set of royalty. Sentence 7c is syntactically different but remains comparable because *dead* and *alive* help to solidify an opposition between the two presidents, *Kennedy* and *Clinton*.

Geographic B-pairs

8a *Munich* was widely hailed as a **success**, *Reykjavik* a **failure**.
8b A separate poll of consumers in the US and Japan showed growing

optimism among *Americans* in contrast to deepening **pessimism** in *the Japanese population*.

8c Historians have largely only differed on whether they saw the German 'takeover' of the empire as a **good** thing (if they were *German* historians) or a **bad** thing (if they were *French or Italian*).

The sentences above use their A-pair to reinforce a contrast between places or nationalities. In sentence 8a, a familiar pair of antonyms (*success/failure*) is helping to create an instantial contrast between two cities (*Munich/Reykjavik*). The B-pair of sentence 8b differs only in that it refers to inhabitants of places rather than the places themselves: *Americans* are contrasted with *the Japanese population*. Finally, sentence 8c differs trivially in that the B-pair here consists of a single nationality (*German*) and a pair of nationalities (*French or Italian*). However, the principle (of a familiar antonymous pair imparting an element of contrast to an instantial opposition) remains the same.

Temporal B-pairs

9a The **bad** news is now *largely behind*, the **good** news is *to come*.

9b The most interesting is that countries which have, in the *Eighties*, done rather **badly** will, in the *Nineties*, do rather **well**.

9c What was **immoral** and **unnecessary** *six months ago* cannot be **moral** and **necessary** *today*.

The B-pair in each of the above sentences concerns the passage of time. For example, sentence 9a contrasts the past (lexically identified as *largely behind*) with the future (*to come*). This temporal contrast is well established in our intuition, so one could not say that the antonymous pair (*bad/good*) actually creates this opposition; rather, two individual oppositions unite to enhance the larger contrast of the sentence, that between each of the clauses. Time is reflected more specifically in sentence 9b, where *the Eighties* is contrasted with *the Nineties*, an opposition signalled by *badly/well*. Sentence 9c is more complex because it contains three oppositions, two of which are morphological antonymous pairs (*immoral/moral; unnecessary/necessary*), and one of which is a temporal contrast between *six months ago* and *today*. Having two established antonymous pairs rather than one makes little difference to this sentence – *immoral and unnecessary* works very much as a single unit which is contrasted with *moral and necessary*. The temporal contrast is less fixed (*six months ago* versus *today*), but is given significant contrastive weight by its antonym-laden context.

Quantitative B-pairs

10a The other aspect of his plan is that his Bill would end the *two*-year wait for **uncontested** divorces and *five* years for **contested** ones.

10b On the question of extending the embargo to cover food and medical supplies, *40* per cent **agree** but *45* per cent **disagree**.
10c It was the old story: **success** has *many* fathers, **failure** has *none.*

In each of the examples above, the A-pair signals a contrast between a numeric (or quantitative) value. For example, in sentence 10a, the numbers *two* and *five* work alongside the A-pair *uncontested* and *contested* to create a dual contrast. The more established opposition is between *uncontested* and *contested*, but numbers themselves are co-hyponyms which seem to have an inherent potential for contrast, especially in the case of low figures such as *two* and *five*. Sentence 10b reports the results of a survey, contrasting the percentage of those who *agree* (40) with the percentage of those who *disagree* (45). Sentence 10c is closer to previous examples of Ancillary Antonymy in that the established opposition (*success*/*failure*) signals that *many* and *none* are the primary contrast. It is interesting to note that the three sentences examined in this chapter which feature numeric values as their B-pair (sentences 5d, 10a and 10b) all 'front' this information by placing it before the A-pair. This is unusual (because the A-pair appears first in the majority of Ancillary Antonymy sentences) and raises the possibility that the numeric contrast is considered to be the less important opposition in such sentences.

Synonymous B-pairs

11a Then, and now, the Royal Festival Hall is a cool, rather clinical building that it is **easy** to *respect* and **difficult** to *love*.
11b Archer was a formal, eccentric man, **long** on *acquaintances* and **short** on *friends*.
11c The West German authorities demurred: under West German law creme de cassis had too **low** an alcohol content to be classed as a *liqueur* but too **high** an alcohol content to be considered a *wine.*

Near-synonyms[2] are not the sort of words one would expect to constitute a B-pair, but the A-pair in each of the sentences above is helping to create a contrast between two words which themselves operate along a scale and occupy similar semantic points on that scale. For example, the B-pair of sentence 11a consists of the non-finite verbs *respect* and *love*. These words could be said to operate along the scale of affection, where they maintain a relatively close proximity to one another. And as one would not ordinarily expect these words to function contrastively, their opposition must be created by context. Given the absence of a contrastive conjunct such as *but*, it would seem that this opposition is licensed predominantly by the A-pair, *easy*/*difficult*. Sentence 11b uses the A-pair *long*/*short* (employed here in a metaphoric sense, as is the norm for this pair) to help create a contrast between *acquaintances* and *friends*. Once again, these B-pair terms would usually be regarded as near-synonyms, but here the A-pair forces the reader to interpret *acquaintances* and *friends* in

terms of their latent contrastive potential. A third example is sentence 11c which uses an A-pair (*high*/*low*) to exploit a potential distinction between a *liqueur* and a *wine*.

Meronymous B-pairs

12a The day's business opened with other foreign matters, Foreign Office questions, during which Secretary of State Douglas Hurd said Britain would welcome a return by South Africa to the Commonwealth as 'a **happy** *end* to a **sad** *chapter*'.

12b But a couple of Libyans are only likely to be **small** *minnows* in a very **large** *pond*.

12c But a Romanian dissident recently dismissed the new regime as 'the same **old** *brothel* with **new** *whores*'.

A relationship similar to that of meronymy[3] is reflected by B-pairs in each of the examples above. The 'part-whole' aspect of sentence 12a concerns the relation between *end* and *chapter*, which, even if considered metaphorically, evidences meronymy. Sentence 12b uses its A-pair (*large*/*small*) to draw out a contrast between *minnows* and *pond*, and sentence 12c uses its A-pair (*old*/*new*) to draw out a contrast between *whores* and *brothel.* Both of these examples refer to a sense relation that could, in general, be termed 'part-whole', or, more specifically, 'inhabitant-location'. The majority of Ancillary Antonymy sentences which feature a meronymous B-pair make use of metaphoric concepts.

Linguistic B-pairs

13a In this account, the **rich** get *to choose*, and the **poor** get *the queues*.

13b For at least one viewer, who had regarded **male** wrestlers as *morons* and **female** wrestlers as *oxymorons*, it was an enlightening experience.

13c Baxter's **active** *can-do* has been overtaken by the **passive** *why-bother*.

The B-pair terms in the sentences above are linked not by any semantic properties, but by the phonetic, morphological and visual constitution of the words themselves. For example, though *to choose* and *the queues* are consequentially related in sentence 13a, their major similarity is phonetic; this B-pair is valid because *choose* rhymes with *queues*. Sentence 13b is even more creative. Here the antonyms *male* and *female* mark a contrast between *morons* and *oxymorons*. Whilst there is obviously some semantic justification for the selection of this lexis, one feels that the writer is contrasting these words primarily because they could be (mis)interpreted as being morphologically related. Sentence 13c illustrates that a B-pair can even include artificial

compounds. Here, *active* and *passive* signal a contrast between *can-do* and *why-bother*, two words linked more by their derivation (and indeed their visual appearance) than by their semantic similarity.

Therefore, these B-pairs are a combination of variable and constant: the difference between *choose* and *queues* is a phoneme or two; the *-u:z* part of each word is identical. Likewise, *morons* and *oxymorons* differ only by the addition of *oxy* and have *moron* in common. The similarity between *can-do* and *why-bother* is that two hyphenated word strings (one a verb phrase, the other a self-standing utterance) have been transformed into nouns; the difference lies in their antithetical meaning.

How ancillary are A-pairs?

The heterogeneity of B-pairs is evidenced by the diversity of the categories outlined above. However, A-pairs are equally flexible in text because they are able to serve different roles in different Ancillary Antonymy sentences. This section will investigate just how 'ancillary' ancillary pairs really are. Do A-pairs actually create contrast between B-pairs or do they simply signal an already established opposition?

The B-pair in every Ancillary Antonymy sentence is endowed with greater contrastive power by the A-pair to a greater or lesser degree. However, it is clear that this degree is much greater in some cases and much less in others. This is dependent on the level of inherent contrast which already exists in the B-pair. For example, some antonyms actively create an opposition between a pair of concepts which would not otherwise be considered contrastively at all. Conversely, some antonyms merely help to affirm an opposition between a pair of concepts which already have a well-established contrastive profile. Between these two extremes, the majority of Ancillary Antonymy sentences use an A-pair to draw out a latent contrast between a pair of co-hyponyms (or even latent co-hyponyms). These words may have the potential for contrast, but would usually be regarded as non-contrastive.

The nine batabase sentences recorded below exemplify the extent to which a B-pair's inherent level of contrast may vary. The first triplet of sentences show an A-pair working alongside a B-pair that is very high in contrast; the second triplet of sentences show an A-pair drawing out a contrast between a pair of co-hyponyms; the third triplet of sentences show an A-pair creating an opposition between a pair of words or phrases which would not otherwise be considered contrastive at all.

14a It's certainly rare to hear anyone speaking about the future of British production in terms of its boundless potential; one can only hope that the next few years prove Puttnam's **optimism** *justified* and his **pessimism** *groundless*.

14b As the old adage put it, *oppositions* do not **win** elections; *governments* **lose** them.

14c Now these orders of time have been reversed: the ***rich*** *rise at dawn*; the **poor** *sleep late*.

15a It is at the moment **illegal** to buy *a bible* on Sunday, even from a cathedral shop, but perfectly **legal** to buy *pornographic magazines*.
15b I bicycled to work, as it was the fastest way of getting there, and as a result I was constantly in danger of death by car, either **slowly** from *asphyxiation* or **quickly** from *being run over*.
15c During the Eighties it was **easy** to obtain consent to build *Canary Wharf* and **difficult** to obtain consent to build *an ex-urban house in Wiltshire or Suffolk*.

16a The new edition appeared in the United States about two weeks ago; when I heard the news of the coup it seemed **bad** news for *democracy*, but very **good** news for *the book*.
16b Now it denotes **high** *butter mountains* and a ***low*** *boredom threshold*.
16c Heathcote Williams' Whale Nation (Cape) backed all the **right** *animal* causes but all the **wrong** *poetic* ones.

It could be argued that sentence 14a features two antonymous pairs. However, the words *justified* and *groundless*, while forming a strong semantic opposition, perhaps lack the coreness of antonymy which *optimism* and *pessimism*, for example, seem to possess. Similarly, sentence 14b contrasts *oppositions* with *governments*, a familiar contrast in contemporary culture which draws on the implicit antonymy of the word *opposition*. However, these two words do not yet appear to be fixed in the mental lexicon to the same degree as, say, the other antonymous pair in this example, *win* and *lose*. Finally, sentence 14c illustrates that the same phenomenon may apply to a multi-word expression. Here, *rise at dawn* is pitted against *sleep late*. These expressions are semantically opposed, but the language is yet to offer any lexis which reflects these concepts sufficiently for them to become enshrined as antonyms. Therefore, this first triplet of sentences shows that the B-pair may itself be a pair of antonyms or near-antonyms.[4] In such examples, it would be folly to argue that the more familiar antonymous pair 'creates' a fresh opposition. Rather, it simply works in tandem with the B-pair and contributes to a larger inter-clausal contrast.

More influential ancillary pairs can be seen in sentences 15a–c. The A-pair in each of these sentences plays a fairly active role in drawing out the latent contrastive potential of a co-hyponymous pair. In sentence 15a, the B-pair is *a bible* and *pornographic magazines*. Both are, of course, publications of some description, but here the writer's intention is to accentuate their dissimilarity, rather than their shared status as co-hyponyms. Sentence 15b contrasts *asphyxiation* with *being run over*, two ways to die. The A-pair here (*quickly*/*slowly*) signals that these concepts are to be interpreted contrastively and also signals the scale against which they are to be interpreted

contrastively, namely the speed of death. This co-hyponymous pair, note, have even less contrastive potential than *a bible* and *pornographic magazines*, which, though also co-hyponymous, seem able to contrast along a secondary scale which could perhaps be termed 'morality'. The B-pair of sentence 15c are buildings, namely *Canary Wharf* and *an ex-urban house in Wiltshire or Suffolk*. Although co-hyponyms, the choice of these two particular places seems to reflect irony: they belong to the same scale, but to opposing ends of that scale if one considers size or monetary value. However, the A-pair of this sentence is still necessary to confirm that the two buildings are to be interpreted contrastively in this context. Therefore, the A-pairs in this triplet of sentences can be seen to function differently from the A-pairs in the previous triplet. Co-hyponyms, by definition, contain some contrastive potential but these B-pairs rank much lower down the scale of innate opposition than their counterparts in sentences 14a–c. Hence, the A-pair is more active in generating B-pair contrast here.

The third triplet of sentences are more extreme again. Here, the B-pair appear to hold no contrastive potential, latent or otherwise. For example, *democracy* and *the book* are the B-pair of sentence 16a. These two concepts seem entirely unconnected, although, within their given context, they must be related in some sense (presumably because *the book* is seen as trivial and personal, whereas *democracy* is seen as universally important). This contrast is signalled by the antonyms *good* and *bad*, which force us to treat *democracy* and *the book* as instantial 'opposites' (or at least as instantial co-hyponyms), regardless of their usual status as lexically unrelated words. The A-pair of sentence 16b (*high* and *low*) serve the same function in respect of *butter mountains* and *a boredom threshold*. Though these phrases are faintly related (the word 'Europe' springs immediately to mind), they are hardly co-hyponyms, nor do they belong to any other specific sense relation. Once again, it is their proximity to a familiar antonymous pair which gives these concepts their momentary contrastive power. Finally, the words *animal* and *poetic* hold no obvious relationship, and, even when their noun head (*causes*) is added, the mist is only partially cleared. Indeed, the words which comprise the B-pair of sentence 16c only become contrastive when their A-pair (*right/wrong*) activate this opposition. In short, the antonyms of sentences 16a–c are effectively creating a comparison between a pair of words with very little innate contrastive value.

To summarise, all Ancillary Antonymy sentences contain an A-pair. However, not all A-pairs are alike; some are more ancillary than others. This depends largely on the B-pair of the sentence: if these words are already highly contrastive (the extreme example being when the B-pair are themselves antonyms and thus indistinguishable from the A-pair), the degree to which the antonyms are ancillary is limited; if, conversely, these words share no obvious semantic relationship, the A-pair is extremely active in creating an opposition between otherwise unopposed concepts. In text, the majority of B-pairs fall somewhere between these two stools and their corresponding A-pairs are thus used to establish an instantial contrast between a pair of co-hyponyms.

Parallelism

Aside from the A-pair, the most important contrast-generating element of an Ancillary Antonymy sentence is parallelism. Typically, two clauses, each following a similar grammatical construction and containing one A-pair member and one B-pair member, will be presented in opposition. Though adversative conjunctions are usually regarded as being most instrumental in creating inter-clausal contrast,[5] the sentences below indicate that textual parallelism, together with an active A-pair, is often sufficient to generate B-pair contrast without the aid of any conjunction.

17a There is *praise* for **success**, *condemnation* for **failure**.
17b The *peace* is usually **male**, the *disturbance* **female**, though in two stories the positions are reversed, and one story, The Image Trade, dispenses altogether with the tension of gender.
17c He leans forward and quotes from a piece of writing in French by Samuel Ullman, which roughly translates as: 'You are as **young** as your *faith*, as **old** as your *doubts*.'

The conjunction in each of the above examples is conspicuous by its absence. For example, sentence 17a uses its A-pair (*success*/*failure*) to signal an opposition between a pair of words which are also contrastive, though perhaps less enshrined in language (*praise*/*condemnation*). Given the high level of parallelism between the two clauses, a further contrastive signal in the form of a conjunction is considered unnecessary. In other words, the grammatical parallelism between *praise for success* and *condemnation for failure* is sufficient to signal that these clauses are to be interpreted contrastively. Sentence 17b is similar. The clause *the peace is usually male* is here followed by *the disturbance female*, a repetition that requires us to insert *is usually*, but otherwise remains the same as clause 1, except for the substitution of A-pair and B-pair members. In sentence 17c, *young* and *old* mark a contrast between *faith* and *doubts*. However, where one might expect to see a *but* (or at least an *and*), only a comma can be found. This demonstrates that parallelism, when used in Ancillary Antonymy contexts, diminishes the need for a contrastive conjunction, and is an important contrast-generating device.

Sentence 17a

		B-pair		**A-pair**
clause 1:	there is	praise	for	success
clause 2:	[there is]	condemnation	for	failure

Sentence 17b

		B-pair		**A-pair**
clause 1:	the	peace	is usually	male
clause 2:	the	disturbance	[is usually]	female

Sentence 17c

		A-pair		B-pair
clause 1:	you are as	young	as your	faith
clause 2:	[you are] as	old	as your	doubts

Conjunctions

A-pairs and parallelism have been identified as key contrast-generating mechanisms in Ancillary Antonymy sentences, even though the more traditional way of signalling inter-clausal opposition is to use a conjunction. However, this section will demonstrate that conjunctions are actually one of the more dispensable contrast-generating devices. Indeed, Ancillary Antonymy sentences which use the adversative conjunction *but*, such as those recorded below, are comparatively rare.

18a *Bofors* might indicate **failure**, but *Venus and Saturn* spell **success**.
18b The First Division of the Endsleigh League is like a well – **easy** to *fall into* but **difficult** to *get out of*.
18c He was perceived as being **able** to *manoeuvre in a Cold War forum*, but **unable** to *adapt to new realities*.

Sentence 18a can be said to create contrast between its two clauses in at least three different ways: first, as explored, the A-pair (*failure/success*) signal an opposition between the B-pair (*Bofors* and *Venus and Saturn*); second, the two clauses of the sentence parallel strongly with one another (SUBJECT + VERB + OBJECT *but* SUBJECT + VERB + OBJECT where the verb phrases (*might indicate* and *spell*) are synonymous); and, third, the word *but* acts unambiguously as a signal that what comes next should be contrasted with what went previously.

Sentences 18b and 18c are comparable in their usage of *but* as a signal of contrast and in their high level of parallelism: in the former, the clauses *easy to fall into* and *difficult to get out of* both use the construction ADJECTIVE + INFINITIVE VERB + PREPOSITION PHRASE; in the latter, *manoeuvre in a Cold War forum* and *adapt to new realities* both use the construction VERB PHRASE + PREPOSITIONAL PHRASE. Such parallelism, coupled with a strong A-pair, could allow *but* to be omitted entirely from these contexts. Therefore, in addition to signalling contrast, it is possible that *but* may serve another function, perhaps to signal unexpectedness (see Winter 1982: 110). In each of the above examples, what follows *but* could be seen as the 'surprise element' of the sentence. However, the point should be stressed that the majority of Ancillary Antonymy sentences do not use a contrastive conjunction; antonymy alone is usually considered sufficient to signal clausal opposition. Indeed, when a conjunction is used to link clauses, it is more likely to be additive than adversative:

19a International support is **long** on *words* and **short** on *deeds*.
19b You want your friends to **hate** the *sin* and **love** the *sinner*.

19c On Saturday night, as news of Claudio's death spread, the *police presence in Vaulx* was **heavy**, and the *violence* relatively **light**.

In the sentences above, *and* is used to connect a pair of contrastive clauses. For example, sentence 19a relies on the antonymous pair *long/short* to signal a contrast between *words* and *deeds*, two terms with a relatively high sense of inherent contrastive potential. The framework is typical of an Ancillary Antonymy sentence because of its parallelism (*X on x and Y on y*) and its contrastive elements, which are syntactically symmetric. In other words, a single clause (*International support is long on words*) is repeated with ellipsis (of subject and verb) and substitution of A-pair and B-pair (*short* replaces *long*; *deeds* replaces *words*). In employing such familiar signals of contrast, the writer effectively renders *but* redundant. Sentence 19b is similar, both syntactically and in its near-idiomaticity. Here, the repetition of the grammatical structures *hate the sin* and *love the sinner* are sufficient to signal that *sin* and *sinner* should be interpreted contrastively. Sentence 19c is a more complex example but the underlying principle remains the same. Here, the antonyms *heavy* and *light* signal a contrast between the noun phrases *police presence in Vaulx* and *violence*. These clauses mirror one another except for the omission of *was* in the second clause. Once again, the cumulative effect of these signals (antonymy, parallelism and ellipsis) seems to diminish the need for a further signal of opposition in the form of a contrastive conjunction.

A number of database sentences link two clauses together without using either *and* or *but* and the examples below succeed in making their inter-clausal contrast explicit without conforming to a standard Ancillary Antonymy structure.

20a While **success** is *sexy*; **failure** is *on a par with cheesy feet*.
20b While *many* **succeed**, however, *a significant number* **fail**.
20c Not only did the IMF **implicitly** *reject US calls for measures to strengthen growth in the industrial world*, it **explicitly** *dismissed demands for a more expansionary Japanese fiscal policy*.

Sentence 20a uses an A-pair (*success/failure*) to help signal a contrast between the word *sexy* and the phrase *on a par with cheesy feet*. Only a semi-colon lies in between the two clauses, and the opposition is marked overtly by *while* at the start of the sentence. Therefore, the framework can be seen as *while X is x, Y is y*. The corresponding structure of sentence 20b would be *while x X, y Y*, where the antonymous pair (*succeed/fail*) mark a contrast between the quantitative B-pair, *many* and *a significant number*.[6] Sentence 20c uses different lexis to achieve the same results. In this example, the first clause is preceded by *not only did*, which seems to have an effect very similar to the *while* of sentences 20a and 20b. The antonyms *implicitly* and *explicitly* then signal a contrast between two verb phrases and their corresponding objects (*reject US calls for measures to strengthen growth in the industrial world* and *dismissed demands for a more expansionary Japanese fiscal policy*).

One effect of using subordinators in this way is to signal that the first clause contains 'given' information, while that contained in the second clause is 'new'.[7] This is notable in all three examples, especially sentence 20b in which *however* confirms the impression that the success of the many is pre-established information.

A-pair/B-pair detachment

A minority of database sentences belonging to the class of Ancillary Antonymy are unusual in that their A-pair and B-pair, instead of alternating, are sequential. In other words, in the sentences exemplified so far, the sequence of A-pair and B-pair has been ABAB or BABA; in the examples below, the sequence is BBAA or AABB. Contexts in which the A-pair and B-pair are syntactically detached are interesting to explore because they show how the ancillary effect can be manipulated, and B-pair contrast generated in unorthodox fashion.

21a As does the absence of easily identifiable *heroes* and *villains*, characters to **love** and characters to **hate**.

21b Around the cornices of Greek temples (as of the Royal Opera House or of Buckingham Palace) there runs the *egg* and *dart* carving which symbolises the **feminine** and the **masculine** principle.

21c Such divorceless marriages and intractable moral issues are the stuff of Keepers of the Flame, which makes an understandably disenchanted survey of what Henry James brilliantly calls 'the quarrel beside which all others are mild and arrangeable, the eternal dispute between the **public** and the **private**, between *curiosity* and *delicacy*'.

Sentence 21a matches the word *love* with the word *heroes* and matches the antonym of the former (*hate*) with the near-antonym of the latter (*villains*). The effect is much the same as in other Ancillary Antonymy sentences: each antonym works in unison with one half of the B-pair to create a larger contrast, in this case between *heroes . . . characters to love* and *villains . . . characters to hate*. Similarly, sentence 21b uses the A-pair *feminine/masculine* alongside the B-pair *egg/dart*, and sentence 21c uses the A-pair *public/private* alongside the B-pair *curiosity/delicacy*. The sequence of A-pair and B-pair in sentences 21a–c is BBAA, BBAA and AABB respectively.

Therefore, so strong is the contrast generated by antonyms that, on occasions, writers can afford to place the B-pair elsewhere in the sentence, rather than interweaving it with the A-pair. This results in two kinds of antonymous usage being exploited: the inclusive effect of coordinating an established antonymous pair (to be discussed in the next chapter) and the ancillary effect of using an A-pair to impart additional contrastive power upon a B-pair.

The ancillary effect

If a sentence features a recognised antonymous pair, the chance of it also featuring a second, related contrast is as high as 40 per cent. In such sentences, the familiar opposition is used to signal a more important contrast between a pair of words (often co-hyponyms) which have less inherent dissimilarity. This phenomenon has been labelled Ancillary Antonymy.

In the database sentences analysed, antonymy appears to be the most powerful signal of contrast. However, parallelism of structure is also noticeable in almost all examples and this too helps to generate contrast. Traditionally, the adversative conjunction was thought to be a powerful signal of contrast (Nesfield 1898: 79), but corpus evidence suggests that this is perhaps the most dispensable contrast-generating device. Many sentences prefer cumulative conjunctions such as *and*; others use no conjunction at all.

The role of 'opposites' in Ancillary Antonymy sentences is to augment the contrastive power of B-pairs. The range of words and phrases which can act as B-pairs is extensive and eight heterogeneous categories have been outlined, though many more could exist as so few grammatical or semantic restrictions are placed on B-pairs. Some B-pairs feature no innate element of contrast at all, while other B-pairs are almost valid antonyms in their own right. Given that at least one sentence in 500 can be expected to feature Ancillary Antonymy, B-pair analysis is an area which would be particularly receptive to further research.

5 New classes of antonymy II
Coordinated Antonymy

When an antonymous pair in the database was found to signal inclusiveness or exhaustiveness of scale, it was assigned to the class of Coordinated Antonymy. Three examples are recorded below:

22a While pensions will not be abolished, the government will encourage everyone, **rich** and **poor**, to rely for their retirement mainly on money they invest in private pension funds.

22b Today, the pressure to make hay while the sun fitfully shines has led to a massive slump in both **public** and **private** standards.

22c Whether he was **right** or **wrong** to raise a certain matter in the way he did, Mr Lawson offered an important insight into his, and almost certainly Mrs Thatcher's and John Moore's, thinking about the long-term future of the welfare state.

In the first of these sentences, *rich and poor* reaffirms the inclusiveness of *everyone* and identifies the scale (*wealth*) against which this inclusiveness is measured. Similarly, sentence 22b refers to *both public and private* standards, thereby placing the accent on inclusion once again, and sentence 22c dismisses the importance of whether Mr Lawson was *right or wrong* to raise a certain matter. This exhausts the scale in question and, once again, shows antonyms functioning in a 'coordinated' fashion – they work in unison, creating no overt element of contrast, and encompass all points on their given scale.

Distribution of Coordinated Antonymy

Of the 3,000 sentences retrieved from the corpus, 1,151 contain an antonymous pair belonging to the class of Coordinated Antonymy. In other words, approximately 38.4 per cent of sentences which feature an antonymous pair uses that pair to express a quality of inclusiveness or exhaustiveness. All fifty-six antonymous pairs occur in a Coordinated Antonymy sentence on at least one occasion, a fact which further illustrates the pervasiveness of the class. Indeed, many individual word pairs function in a coordinated[1] fashion at a rate similar to the global distribution across the whole database. For example, 39.0

Table 5.1 Antonymous pairs expressing Coordinated Antonymy most commonly

	Database sentences	*Coordinated Antonymy attribution (%)*	*Normal distribution score*
confirm/deny	34	**100.0**	1.00
disprove/prove	14	**100.0**	1.00
rightly/wrongly	44	**97.7**	1.00
directly/indirectly	79	**72.2**	0.94
explicitly/implicitly	30	**63.3**	0.87
guilt/innocence	44	**61.4**	0.85
correct/incorrect	18	**61.1**	0.85
gay/straight	33	**60.6**	0.85
false/true	62	**54.8**	0.77
illegal/legal	31	**54.8**	0.77

per cent of *cold*/*hot* sentences have been attributed to this class, as have 37.8 per cent of *happy*/*sad* sentences and 38.9 per cent of *advantage*/*disadvantage* sentences. Remarkably, 38.4 per cent of the 156 *un*-word sentences function in terms of Coordinated Antonymy, exactly the same proportion as identified across the entire 3,000 sentence database.

The ten antonymous pairs which are coordinated most frequently are presented in Table 5.1: the figures in the first column record how many sentences featuring each pair were included in the database; the percentages recorded in the second column refer to the proportion of those sentences attributed to Coordinated Antonymy; and the normal distribution score in the third column shows how strongly, in relative terms, each pair favour this class.[2]

Table 5.1 shows that none of the thirty-four *confirm*/*deny* database sentences feature their antonymous pair functioning in anything other than a coordinated environment. Similarly, all fourteen *disprove*/*prove* sentences belong to Coordinated Antonymy, as do all but one of the forty-four *rightly*/*wrongly* database sentences. Indeed, this trio of word pairs all achieve a normal distribution score of 1.00, which indicates that their bias towards Coordinated Antonymy cannot be attributed to chance. Therefore, usage such as that reflected by the three sentences below is common to the point of being inevitable.

23a Whitehall was yesterday unable to **confirm** or **deny** other simulated devolutions.

23b What we can't even begin to do is evaluate any mystical or imaginative statement – we cannot **prove** or **disprove** the idea that God is angry.

23c We helped a landbound frog, **rightly** or **wrongly**, back to the water's edge and pedalled on to collect our reward: two Tangle Twisters from the kiosk by the boat-hire, and time to sit in the English afternoon sun and watch the world drift by.

The sentences above are typical of those favoured by this trio of antonymous pairs. The negativity of sentences 23a and 23b is also in keeping with this type of context: 'Whitehall is *unable* to confirm or deny' a proposition and 'we *cannot* prove or disprove' a given idea. *Rightly/wrongly* does not function in a coordinated fashion throughout the database, but can be seen to favour this framework in a large majority of sentences. It would appear that these three antonymous pairs have been used in these sorts of contexts with such frequency that they now border on idiomaticity; to use them in any other way almost feels wrong.

Surprisingly, Table 5.1 reveals that none of the six pairs which show the greatest bias towards Coordinated Antonymy are adjectival. This is particularly striking given that over half of the antonyms selected for study are adjectives. Both of the pairs that score 100 per cent are verbs, the next three highest-scoring pairs are adverbs, and the sixth-placed pair are nouns. Only then do adjectival antonymous pairs such as *correct/incorrect* and *gay/straight* arise.[3]

Frameworks of Coordinated Antonymy

Coordinated antonyms are usually conjoined by *and* or *or*. In general terms, those linked by *and* can be seen as 'inclusive'; those linked by *or* as 'exhaustive'. Neither sub-class features any element of distinction between the two antonyms – in both cases, the antonymous pair is presented equally by the text. In other words (and perhaps paradoxically), it is the similarity between the antonyms (i.e. their shared status as co-hyponyms of a given super-ordinate), rather than their inherent semantic dissimilarity, which is the primary focus of attention. This reflects Clark's 'minimum contrast rule' (1970: 275), Leech's inclination towards componential analysis (1974: 98) and Cruse's rhetoric about the 'closeness' of antonyms (1986: 197), all of which draw attention to the fact that 'opposites' are word pairs which have much in common.

This section will analyse those pairs which occur in an *X and Y* framework, then analyse those pairs which occur in an *X or Y* framework. Finally, antonymous pairs which occur in neither of these frameworks but still belong to the class of Coordinated Antonymy will be examined.

X and Y

Of the 1,151 database sentences classified in terms of Coordinated Antonymy, 490 make use of the framework *X and Y*. Such contexts use antonymy to signal that both halves of a given semantic scale are applicable and nine database examples are recorded below:

24a In line with a shell and chassis design geared to **active** and **passive** safety, the engines in the low and middle range have been built with economy and pulling power in mind.

24b Both pictures are attributed to Ambrogio Lorenzetti, one of the foremost painters in Siena in the first half of the fourteenth century, who is better known for his massive frescos on **good** and **bad** government in the Palazzo Publico, the city hall.

24c And this fear of going naked into the presence of papa seems particularly apposite when papa is Kenneth Clark, author of The Nude, a book through which many a schoolchild has flicked with just those confusing sensations of **light** and **heavy** tumescence.

25a A frank exposé of the oldest profession – **male** and **female** – in Budapest.

25b The idea of poets **alive** and **dead** swapping imagery in order best to express feeling is heartening; since life often seems beguiling and confusing, a little help in capturing the paradoxes is always welcome.

25c Ahnenerbe also, however, administered Lebensborn, a 'welfare facility' which looked after racially 'pure' German mothers, **married** and **unmarried**, during their pregnancy and took over many of their children when they were born.

26a Again in debates over genetic research it is significant that Christians, Muslims and Jews have united, **implicitly** and **explicitly**, in condemning a low view of the value of embryonic life.

26b He took **success** and **failure** in his stride.

26c Editors have moved at roughly 10-yearly intervals and steered the journalists towards the provocative leader-articles that The Economist's readers **love** and **hate**.

The first triplet of sentences above show an antonymous pair directly pre-modifying a noun phrase (*X and Y n*). Such antonymy is contextually removable from the sentence; its function is to signal inclusiveness. Thus, we discover that the safety of sentence 24a is both *active and passive*, the governments of sentence 24b are both *good and bad* and the tumescence of sentence 24c is both *light and heavy*. The effect of these antonymous phrases is to inform us that the subsequent noun head refers to both ends (and usually everything in between) of the antonymous scale against which it is being measured.

Sentences 25a–c are semantically similar, but here the antonymous pair post-modify the noun head (*n*, *X and Y*) rather than pre-modify (*X and Y n*). Primarily, this is a rhetorical device, although *male and female* could not act as pre-modification in sentence 25a because the noun head is implied rather than stated (*members* of the oldest profession). Similarly, *alive and dead* could not act as pre-modification in sentence 25b because *alive* acts only as subject complement. Perhaps the reason why *married and unmarried* post-modifies *mothers* in sentence 25c is that it would give the already full pre-modification a 'cluttered' feel. Therefore, it can be stated that an adjectival antonymous pair,

if coordinated by *and*, is more likely to pre-modify than post-modify a noun head, but remains flexible enough to post-modify should more pressing contextual or rhetorical factors demand.

Given that inclusiveness appears to be the primary goal of *X and Y* constructions, one would not expect to find non-gradable pairs used as antonyms. Non-gradable pairs exhaust their entire scale anyway, so what would be the point of signalling inclusiveness?

If signalling inclusiveness was the only function of Coordinated Antonymy, then the antonymous pair of sentences 25a–c would be redundant. After all, people can only be *male* or *female*, poets can only be *alive* or *dead* and mothers can only be *married* or *unmarried*. However, another function is being served by these antonyms: they defy our expectations. In other words, the reason that inclusiveness needs to be signalled in these sentences lies in the nature of the noun head being modified. For example, without the phrase *male and female*, one would be inclined to assume that the members of the oldest profession would be exclusively female. Thus, the antonymous pair is defying our preconceptions and the phrase can be seen as being weighted more towards *male* than *female*. The writer is effectively anticipating our assumptions about a concept, then correcting this assumption. So, though *male* is given no overt linguistic priority in the sentence, it is clearly the antonym which carries the greater semantic significance. One could regard it as the 'surprise antonym'.

Sentence 25b is similar. Here, one would assume the poets in question to be *alive* were it not for the antonymous phrase *alive and dead*. Sentence 25c is more difficult to interpret, but given that Lebensborn 'took over' many new-born children, one might expect the mothers in question all to be *unmarried* were it not for the phrase *married and unmarried*. Thus, the antonymous pair of Coordinated Antonymy sentences can be said to fulfil two related functions: to signal inclusiveness of scale and, sometimes, to defy our expectations about a given noun head.

The antonymous pairs featured in the first two triplets of sentences are all adjectival. Sentences 26a–c illustrate that the same framework (*X and Y)* may be occupied by adverbs, nouns and verbs respectively. Instead of antonymous adjectives modifying a given noun head, sentence 26a features antonymous adverbs (*implicitly*/*explicitly*) modifying a given non-finite verb (*united*). The effect appears to be much the same – the antonyms add an encompassing element to the word described, signalling inclusiveness of scale. Being nouns themselves, the antonymous pair of sentence 26b do not act as modification. However, the phrase *success and failure* is also inclusive in this context: *he took success and failure in his stride* is analogous with the idiom 'he took everything in his stride'. Finally, sentence 26c illustrates that it is possible for inclusive coordinated antonyms to take the form of verbs. Here, the antonymous pair (*love* and *hate*) belong to the subject *The Economist's readers*. Once again, the function of antonymy in this sentence is to signal inclusiveness, or, more specifically, inclusiveness which might not otherwise have been expected (*hate* seems to be the 'surprise antonym' here).

X or Y

Of the 1,151 database sentences classified in terms of Coordinated Antonymy, 533 occur in the framework *X or Y*. Often, the semantic difference between *X or Y* and *X and Y* is negligible. Both frameworks serve to encompass a semantic range, but whereas *X and Y* may simply refer to each antonym without necessarily accounting for all in between, *X or Y* tends to symbolise an entire range. Also, antonyms joined by *or* (in contrast to those joined by *and*) are more likely to post-modify than pre-modify their noun head. Six typical sentences are recorded below:

27a Most Ugandans, **married** or **unmarried**, had several lovers.
27b But assuming no scandals, **old** or **new**, precipitate presidential disgrace, what is he to do if a triumphant place in history is to be assured?
27c When governments realise that transport planning means taking a comprehensive view of all transport systems and co-ordinating them, they will find themselves better prepared not only for the country's future, but for making their case against lobbying groups, **large** or **small**.

28a He showed no disloyalty, **publicly** or **privately**, to Virginia Bottomley though it must have irked him that she was in the Cabinet and he was not.
28b The Koran unequivocally states that the **punishment** or **reward** for insulting God lies with Him alone – muslims, mullahs and the courts have nothing to do with it.
28c Yet, **win** or **lose**, he could fade faster than Donny Osmond if the money goes to his head.

In sentence 27a, the antonymous phrase *married or unmarried* expresses exhaustiveness. Most Ugandans, regardless of their marital status, are said to have had several lovers. Sentence 27b is very similar in terms of syntax (*n*, *X or Y*), but is different in the sense that its antonymous pair (*old*/*new*) is gradable, whereas the antonymous pair of sentence 27a is non-gradable. In one respect, it is inevitable that a non-gradable pair will signal exhaustiveness in this context because Ugandans, like the rest of us, can only fall into one of two possible categories – *married* or *unmarried*; *old* and *new*, by contrast, have intermediary and extreme stages and do not necessarily exhaust every point on their scale. However, it would seem that the gradable/non-gradable distinction is largely academic when applied to these two sentences; *married or unmarried* is inherently exhaustive, but *old or new* is equally (though not inherently) exhaustive because this antonymous phrase effectively represents all points on the given scale, not just the two specific points mentioned.[4] This encompassing quality of coordinated antonyms is also evident in sentence 27c,

in which the antonymous phrase *large or small*, functions as exhaustively as any non-gradable pair. Here, the preceding noun head is *lobbying groups* and the expression seems synonymous with 'lobbying groups of any size' rather than specifically 'either large or small lobbying groups'. Therefore, regardless of whether they are gradable or non-gradable, antonymous pairs coordinated by *or* are able to express exhaustiveness and symbolise an entire scale.

The antonyms of sentences 27a–c are all adjectival. Sentences 28a–c illustrate that the same phenomenon may be evidenced by adverbs, nouns and verbs respectively. In sentence 28a, *publicly or privately* modifies the verb *showed*. As with adjectival *X or Y* contexts, the two words symbolise all points of the *publicly*/*privately* scale, being synonymous with 'of any kind'. Sentence 28b is somewhat different, as the antonymous pair and the noun phrase are one and the same. However, in many respects, *punishment or reward* still functions exhaustively. The writer believes that the consequence of 'insulting God' is a matter for God alone, and both antonyms are recorded as a means of exhausting the entirety of the scale in question. This makes sentence 28b comparable with other exhaustive sentences, even if its syntactic structure is quite different. Finally, sentence 28c illustrates that verbs can also be used exhaustively. Here, *win or lose* signals that the outcome of a given event is irrelevant to the statement being made.

A variation on this framework occurs when *X or Y* is preceded by *how*. This creates a gauging mechanism against which a given concept can be measured.

29a When the parents return, they shall not make the sitter listen to an account of how **good** or **bad** their evening out was.

29b It joined the committee in arguing that the change would make the new figures no longer comparable with old ones – making changes in how **well** or **badly** the NHS was doing impossible to measure.

29c In the autumn the school's 700 pupils will have a homework diary so that parents will know what homework has been set and will be able to comment on how **easy** or **difficult** their children find the assignments.

The antonymous pair in each sentence above exhausts its given scale, but does so in a different manner from those *X or Y* contexts examined previously. For example, in sentence 29a, the antonymous pair is present because the writer feels that a single antonym would be interpreted as reflecting a bias towards its corresponding end of the scale. In other words, in this instance, to say *how good* the evening was implies that the evening was, indeed, *good* (to a greater or lesser degree); however, to say *how good or bad* the evening was implies no bias either way.[5] Indeed, if any bias were to be identified it would be towards *bad*, because the inclusion of this alternative is marked. Similarly, sentence 29b could rely on *well* or *badly* alone to signal the given scale. The latter would

strongly suggest a negative performance by the NHS; the former would be less marked, but perhaps remains inclined towards a positive outcome; only *how well or badly* allows the writer to identify the scale in question without signalling any unwanted, pre-emptive bias. Sentence 29c further illustrates this phenomenon by referring to *how easy or difficult* homework might be.

To digress momentarily, this construction raises questions about the concept of the 'unmarked antonym' (or the 'impartial antonym', as Cruse (1986: 208) prefers). Palmer notes that the unmarked antonym is the 'only one used simply to ask about or describe the degree of the gradable quality' (1976: 80). In other words, if one asks 'how tall is Jack?', one is not necessarily assuming that Jack is in any way tall. However, if one asks 'how short is Jack?', the presumption is that Jack must indeed be short.

In some respects, sentences 29a–c provide grounds for disputing this logic as they do not rely on one antonym to signal impartiality; rather they use both antonyms in a *how X or Y* framework. This suggests that, contrary to Palmer's assertion, *good* is not always sufficiently unmarked to carry no bias at all. If it were, then *or bad* would be superfluous in sentence 29a. Similarly, if *well* were entirely unmarked then *or badly* would be superfluous in sentence 29b, and if *easy* were entirely unmarked then *or difficult* would be superfluous in sentence 29c.

In isolation, these sentences appear to question the validity of stating that one antonym is truly unmarked. However, further examination of corpus data shows that examples such as those above are relatively rare. For instance, 1,006 corpus sentences refer to *how good* something is, but only 13 refer to *how good or bad* something is. Therefore, fewer than 1.3 per cent of *how-* constructions use both *good* and *bad*; on a large majority of occasions, *good* is considered sufficient. Interestingly, the marked antonym is employed alone in 411 corpus sentences (*how bad*), but *how bad or good* occurs just once. It is difficult to imagine neutrality of scale being the objective of *how bad* constructions.

A similar pattern emerges for *well* and *badly*. There are 1,466 corpus sentences which refer to *how well*, but only 18 refer to *how well or badly*. Therefore, writers only consider it necessary to append *well* with *or badly* in only 1.2 per cent of sentences. However, no similar experiment could be conducted for *easy*/*difficult* because sentence 29c records the only occurrence of *how easy or difficult* in the corpus. *How easy* features 384 times and *how difficult* appears 619 times.

What this indicates is that the view espoused by Palmer and others about 'unmarked' antonymy is valid in about 99 per cent of cases when neutrality of scale is to be signalled. However, as the triplet of sentences above show, writers occasionally feel it necessary to include the marked antonym because its partner is still felt to carry some degree of bias. This is perhaps analogous with the increasing popularity of the gender-neutral phrase *he or she* (or *(s)he*), which suggests that writers feel uneasy about using *he* alone when referring to unspecified individuals.

Non-standard frameworks of Coordinated Antonymy

The framework *X and Y* is used in 42.6 per cent of Coordinated Antonymy sentences and the framework *X or Y* is used in 46.3 per cent of Coordinated Antonymy sentences. This leaves 11.1 per cent of Coordinated Antonymy sentences which adhere to neither framework. Some of these less typical contexts feature a framework such as *neither X nor Y* or *X as well as Y*; others feature more unusual frameworks which still reflect Coordinated Antonymy but do so in an idiosyncratic fashion. Before these frameworks are discussed, nine examples are recorded of sentences which do not adhere to an *X and/or Y* framework simply because they have additional lexis which complicate that construction.

30a He played numerous cameo roles both on the **large** and the **small** screen.
30b Only he could have said whether this amounts to a **reward** or a **punishment**.
30c George became careful about touching his daughters, both in **public** and in **private**.

31a 'I've had **difficult** matches and **easy** matches with Mats, and I think this one will be tougher than the Davis Cup,' he said.
31b You soon learn that there are **good** people and **bad**, just like on the factory floor.
31c In our time people either **love** herb gardens or **hate** them.

32a It's a **strength** and, obviously, a **weakness**.
32b But it was not clear whether the ceasefire was intended to be **permanent** or merely a **temporary** measure to allow the evacuation of civilians and wounded.
32c What is needed is a system 'which leaves individuals free to **succeed** or, just as important, to **fail**'.

The function of antonyms in the sentences above does not differ greatly from examples already discussed, but these contexts show how writers vary antonymous frameworks for rhetorical effect. These variations have repercussions for antonym retrieval (see Chapter 10), because standard constructions are manipulated; the reasons for this manipulation will now be discussed.

Sentences 30a–c are notable for an absence of ellipsis: in sentence 30a, *the* is omissible (*the large and {the} small screen*); in sentence 30b, *a* is omissible (*a reward or {a} punishment*); and in sentence 30c, *in* is omissible (*in public and {in} private*). Why writers choose to include these small words when sentences would be no less grammatical in their absence can only be speculated about. In sentence 30b, *a punishment* may have been preferred to *punishment* because the latter could be read as a mass noun if one did not realise that ellipsis of the

indefinite article had taken place. However, no such disambiguating function can be identified in sentences 30a or 30c. Here, lexical efficiency seems to be compromised for stylistic effect. Perhaps, subconsciously, the writer feels that the tone units of these sentences are prosodically preferable with an additional syllable. Arguably, the extra repetition within *the large and the small screen* and *in public and in private* is more rhythmic to the ear. An alternative explanation is that *the small screen* and *in private* are slightly idiomatic; ellipsis of *the* and *in* may be resisted because *small screen* and *private* seem collocationally stranded without these words.

Sentences 31a–c demonstrate the mobility of the noun head in coordinated constructions. In sentence 31a, the noun head (*matches*) is repeated after both antonymous adjectives; in sentence 31b, the noun head (*people*) occurs after the first antonym where one would usually expect it to occur after the second; and in sentence 31c, the noun head (*herb gardens*) again follows the first antonym, while a corresponding pronoun (*them*) follows the second. In each of these three sentences, the standard construction would be *X or/and Y n*. Once again, prosodic factors may help explain their digression from the norm, though one could also argue that specifying the noun head after the first antonym is beneficial to the reader because s/he is not kept in suspense until after the second antonym to discover what is actually being described.

Sentences 32a–c are different again, illustrating that the standard frameworks for Coordinated Antonymy may be interrupted by an adverb or other interpolation. For example, *obviously* and *merely* complicate the frameworks of sentences 32a and 32b respectively, while the antonymous pair of sentence 32c is split by *just as important*. Such additions make the sentence syntactically different from the majority of coordinated sentences, but these examples are rendered no less inclusive or exhaustive for their additional vocabulary.

One might counter-argue a case for this final triplet of sentences to be classified as Ancillary Antonymy if one claimed that a second contrast arises in each. For example, in sentence 32a, the two forms of the verb *to be* are different: *is* and *{is} obviously*. Likewise, in sentence 32b, *to be* contrasts with *{to be} merely*. This could make the contexts analogous with sentences analysed in the previous chapter if these 'contrasts' were identified as being B-pairs. However, it seems far more natural to consider these three sentences as being Coordinated Antonymy laced with modality. The modality itself does not constitute a second contrast. Nevertheless, this minor ambiguity demonstrates that the classification process sometimes requires subjective choices to be made.

neither X nor Y

So far, those sentences examined in this section have avoided conforming to typical frameworks of Coordinated Antonymy by shunning ellipsis or rearranging their syntactic elements. However, other non-standard Coordinated Antonymy sentences use more predictable constructions. For example,

twenty-three database examples feature the framework *neither X nor Y*, three of which are recorded below:

33a Bobby Gould, the Albion manager, would neither **confirm** nor **deny** Goodman's impending departure, saying pointedly: 'You'll have to talk to the directors about that.'
33b If a school with bad exam results says it is, nevertheless, producing fine people, we can neither **agree** nor **disagree**.
33c Thompson, in an interview on BBC Television's Sportsnight programme, says he is neither **pessimistic** nor **optimistic** about his prospects for the Games but is convinced he can still be an athletics force for the next four or five years.

The trio of sentences above negate an antonymous pair in order to signal neutrality. As such, they are not exhaustive. Indeed, they are only inclusive in the sense that they couple together two antonyms in order to negate the pair. Ten of the twenty-three *neither X nor Y* contexts retrieved feature *confirm*/*deny* in X and Y position. A cliché among politicians and celebrities not wishing to answer questions, this expression is approaching idiomatic status. Sentence 33b is very similar to sentence 33a in that a pair of verbs (*agree*/*disagree*) are negated to intimate neutrality of scale. Sentence 33c demonstrates that adjectives can also function in this way.

Once again, one could argue that these sentences do not rightfully belong to the class of Coordinated Antonymy. Here, a case could be made that these sentences express Extreme Antonymy,[6] in which a pair of antonyms are coordinated, but express either end of their scale without exhausting all semantic space in between. However, the parallels between these examples and standard Coordinated Antonymy sentences remain strong – no overt contrast is generated and the antonymous terms work in harmony to identify a given scale.

X as well as Y

Another framework associated with Coordinated Antonymy is *X as well as Y*. Of the 1,151 database sentences attributed to the class, twenty-one were found to use this construction.

34a A sign, perhaps, that **public** as well as **private** allegiance is transferring itself from God to Mammon.
34b It would be interesting to hear all experiences, **good** as well as **bad**.
34c Cantona postured, Chris Eubank-like, relishing the moment, and another few thousand learned to **hate** as well as **love** him.

These sentences are assigned to the class of Coordinated Antonymy because

the phrase *as well as* acts much the same as *and*, giving the antonyms a sense of inclusiveness. As with many inclusive coordinated sentences, the antonymous pair is weighted unevenly. In the above examples, it is the first-mentioned antonym which carries the greater weight and is presented as the 'surprise antonym'. In other words, were it not for the antonymous phrase, one might expect the allegiance of sentence 34a to be *private*; the experience of sentence 34b to be *bad*; and the feelings towards Cantona of sentence 34c primarily of *love*. Similarity between these sentences and standard inclusive Coordinated Antonymy sentences (*X and Y*) is evidenced further by the fact that the antonymous phrase may pre-modify the noun phrase (sentence 34a), post-modify the noun phrase (sentence 34b), or occur in verb form (sentence 34c).

Residual frameworks

The class of Coordinated Antonymy is home to a further thirty-three database sentences which conform to no obvious syntactic pattern. Six are recorded below:

35a Geoffrey Dear, the Chief Constable, purged the lot – the **good** with the **bad** – and scattered them to the furthest reaches of the force's territory.
35b Of the buyout he said: 'We may **succeed**, we may **fail** – but we will at least give it a whirl.'
35c Anyone – **male**, **female**, **old**, **young**, 'civilised', 'primitive' – can become de-individuated.

36a The dark was nurturing; it was where, in church, I was connected to everyone else: **living**, **dead**, present or not, mentally disturbed, outcast, **old**, **young**, **poor**, **rich**, intelligent, of the establishment, or criminal – in fact, everyone gathered around that table.
36b Paramount, he claims, must know it goes on ('They don't **encourage** it, but they don't **discourage** it either').
36c If it's **wet** we can play it up front, if it's **dry** we can play it up front, if it's **wet** we can play it wide and if it's **dry** we can play it wide.

The examples above show how the rhetorical effect of Coordinated Antonymy can be achieved by writers in unorthodox and creative ways, without making use of standard frameworks associated with this class (namely *X and/or Y*). For example, sentence 35a features the structure *the X with the Y*. This is comparable with *X as well as Y* in that the X-position antonym is presented as the 'surprise antonym'.[7] Sentence 35b is notable for an absence of a coordinator between the clauses *we may succeed* and *we may fail*.[8] This prevents the ellipsis of subject and auxiliary verb (*we may*) that one would expect in this context. Sentence 35c is also *or*-less. Here, two established antonymous pairs (*male/female* and *old/young*) and one less enshrined pair of 'opposites' (*civilised/primitive*) are listed as post-modification of *anyone*. Once again, these antonyms

are acting exhaustively, representing the entire range of their given scales, namely gender (a two-point scale) and age (a many-point scale).

Sentence 36a is similar, but makes use of an even longer list. The antonymous pairs *living*/*dead*, *old*/*young* and *poor*/*rich* occur in this list, as do terms representing other groups of people. An interesting coordinated function is served here by the phrase *present or not*; in such constructions, *not* substitutes for an antonym[9] and generates a very strong sense of exhaustiveness. Sentence 36b is a good example of *but* functioning non-contrastively. This context is reminiscent of *neither X nor Y* sentences in that both antonyms are negated, but here *they don't encourage it* and *they don't discourage it* are conjoined by *but* where one might expect *and*. Finally, sentence 36c is arguably the most complete example of Coordinated Antonymy in the database. Without conforming to any typical syntactic patterns, *wet* and *dry* function exhaustively on two separate occasions. First, they signal that weather conditions are irrelevant in determining whether or not they can *play it up front*; then they signal that weather conditions are irrelevant in determining whether or not they can *play it wide*. This sentence, like many unconventional Coordinated Antonymy examples, is extended by a want of ellipsis. With further compression, sentence 36c might well read: 'wet or dry, we can play it up front; wet or dry, we can play it wide'. And with even more compression, the sentence could read: 'wet or dry, we can play it up front or wide'. This is more syntactically efficient, perhaps, but far less rhetorically effective and prosodically pleasing.

Having examined all of the major frameworks which house Coordinated Antonymy (and a number of minor frameworks), it is now possible to identify lexical signals associated with this class. Framework productivity is investigated fully in Chapter 10, but Coordinated Antonymy is important in this respect because its frameworks are among the most robust of any class.

Productivity of Coordinated Antonymy

Coordinated antonyms are relatively stable in the nature of their syntactic co-occurrence. Of the 1,151 Coordinated Antonymy sentences retrieved, 88.9 per cent of antonyms are positioned two words away from one another in text and conjoined by either *and* or *or*. The majority of these sentences simply refer to *X or Y* or *X and Y*, but others use nearby lexis as signals. These frameworks (and their corresponding frequency in the database) are listed below, together with an illustrative example:

both X and Y (77)

> And this good practice can be employed in respect of **both listed and unlisted** properties.

either X or Y (31)

> Was the course of justice perverted, **either consciously or unconsciously**, by the damage limitation exercise NatWest put in operation in the wake of the affair?

neither X nor Y (23)

Mr Pierson said he was '**neither optimistic nor pessimistic**' about reaching agreements.

whether X or Y (17)

About 52 per cent of women in the UK are economically active – meaning they are considered to be part of the working population, **whether employed or unemployed**.

how X or Y (10)

It is **how well or badly** a person plays a game, runs a race, or rides a horse that matters most.

X and Y alike (9)

These qualities all made him sought after by **young and old alike**.

The most reliable signal is *both*, which immediately pre-modifies *X and Y* in one-fifteenth of all database sentences assigned to Coordinated Antonymy. The framework *either X or Y* arises on thirty-one occasions, accounting for 2.7 per cent of all coordinated sentences; *neither X nor Y* is employed in a further 2 per cent of contexts, and *whether X or Y*, *how X or Y* and *X and Y alike* each account for about 1 per cent of sentences assigned to the class of Coordinated Antonymy. The last-mentioned of these frameworks is the only one which shows a strong bias towards particular antonymous pairs – of the nine *X and Y alike* sentences retrieved, six feature *old*/*young* and three feature *gay*/*straight*.

Though the number of database sentences which correspond to each productive framework may initially seem small, the six word strings above feature collectively in nearly 15 per cent of all Coordinated Antonymy sentences, suggesting that the signals identified (*both*, *either*, *neither*, *whether*, *how* and *alike*) are very significant markers of antonymy in text.

The coordinated effect

Database sentences classified in terms of Coordinated Antonymy signal inclusiveness or exhaustiveness of scale, usually by conjoining antonyms with *and* or *or*. As all fifty-six word pairs under scrutiny function in this way at least once in the database, Coordinated Antonymy should be recognised as a relatively widespread phenomenon in language. On average, antonymous pairs are coordinated in about 38 per cent of contexts; some pairs (*confirm*/*deny* and *disprove*/*prove*) behave this way exclusively throughout the database.

So, whereas Ancillary Antonymy sentences have been shown to maximise the contrastive power of antonyms, Coordinated Antonymy sentences allow that contrastive power to remain dormant. In other words, the former class is home to 'opposites' which disseminate their antonymity; the latter class is home to 'opposites' which discard their antonymity and exploit their latent similarities instead.

6 New classes of antonymy III
Minor classes

The previous two chapters have dealt with the most common functions of antonymy in text, namely to assign additional contrastive value to a nearby pair of words or phrases (Ancillary Antonymy) and to signal exhaustiveness or inclusiveness of scale (Coordinated Antonymy). Over 77 per cent of all database sentences fall into one or other of these two classes. This chapter will discuss the minority of antonyms which are neither coordinated nor ancillary, most of which were attributed to one of six minor classes of antonymy. These classes are recorded in Table 6.1, together with the number of database sentences assigned to each, expressed as raw frequency and as a proportion of the whole database.

The distribution of sentences among minor classes of antonymy is not uniform. Comparative Antonymy is the largest of the minor classes, accounting for 205 database sentences. Distinguished Antonymy accounts for a further 161 sentences, but no other class contains more than 100 sentences. Indeed, the eighth-largest class (Idiomatic Antonymy) contains only twenty-three database sentences, which illustrates how sharply the frequency of these classes decreases. Aside from Ancillary Antonymy and Coordinated Antonymy, all functions of antonymy are relatively low frequency and half of the eight new classes identified feature 3 per cent of database sentences or less. A taxonomy of all six minor classes will now be presented. Although the relative

Table 6.1 Frequency of minor classes of antonymy in the database

Minor class	*Number of database sentences*	*Proportion of database sentences (%)*
Comparative Antonymy	205	6.8
Distinguished Antonymy	161	5.4
Transitional Antonymy	90	3.0
Negated Antonymy	62	2.1
Extreme Antonymy	40	1.3
Idiomatic Antonymy	23	0.8
Residual sentences	*106*	*3.5*
Total	**687**	**22.9**

size of these classes is small, the ways in which antonymy functions at lower levels of frequency demonstrate the flexibility of 'opposites' in text. Only 106 database sentences could not be attributed to any new class of antonymy; these residual examples will be discussed towards the end of this chapter.

Comparative Antonymy

Definition:	The co-occurrence of an antonymous pair within a framework that places those words in a comparative context or measures one antonym against the other.
Typical frameworks:	*more X than Y* *X is more* [adj] *than Y* *X rather than Y*

Only 6.8 per cent of database sentences have been classified in terms of Comparative Antonymy. This infrequency is surprising because the measuring of one antonym against another is among the functions one might intuitively associate with antonymy. Six database examples are recorded below:

37a However, **light** crude is more easily broken down than **heavy** crude from the Middle East, making it less damaging environmentally.
37b The question is perhaps easier to answer for the **long** term than the **short**.
37c Sometimes I feel more **masculine** than **feminine** and I don't like it.

38a Mr Shevardnadze stressed that work for **peace** rather than **war** should prevail.
38b It seems that where **gay** relationships are successful, they are being even more successful than **straight** ones, because people have to work so much harder on what relationships are about.
38c It applies to shareholdings in **large** companies just as much as it does to those in **small** ones.

The first triplet of examples feature comparisons typical of those assigned to the class of Comparative Antonymy: in sentence 37a, *light crude* is compared to *heavy crude* in terms of the ease with which it may be broken down; sentence 37b compares *the long term* to *the short*, in terms of answering a given question; and sentence 37c describes somebody as feeling *more masculine than feminine.*

The second triplet of sentences are slightly different: sentence 38a uses the formula *X rather than Y* in its comparison of *peace* and *war*; sentence 38b compares *gay relationships* to *straight ones*, introducing a pronoun (*they*) in the place of the former; and sentence 38c is unusual in that, unlike the previous five, it does not give priority to one antonym over the other. Rather, *large companies* and *small ones* are compared, but presented as being equal.

The subtle differences arising between instances of Comparative Antonymy

(such as those intimated above) will now be discussed in closer detail. Based on some of these distinctions, the class can be broken down further into four sub-classes. First, Direct Comparison describes contexts in which something is described as containing (or being) *more X than Y*. This is exemplified by sentence 37c. Second, Indirect Comparison describes contexts in which two antonymous concepts are compared with one another in terms of another scale. This is exemplified by sentence 37a in which *light crude* and *heavy crude* are compared to one another along the scale of how easily they may be broken down. Third, Preferential Comparison refers to a framework such as *X rather than Y* in which one antonym is given preference over the other. This is exemplified by sentence 38a. Finally, Equal Comparison describes situations in which two antonymous concepts are compared with one another, but no difference found. This is exemplified by sentence 37c, although it should be noted that Equal Comparison is essentially a sub-class of Direct and Indirect Comparison.

Direct Comparison

39a Although one has to be more **pessimistic** than **optimistic** in the Third World when it comes to liberation movements and democracy, it is hard not to notice some special characteristics of the Palestinian revolution in general and the PLO in particular.

39b And it is possible to accept both that Dr Higgs was a lot more **right** than **wrong** in her diagnoses, but that it is now impossible for her to return.

39c 'Well,' said Cage, completely unabashed, 'some living composers are more **dead** than **alive**'.

The trio of sentences above all conform to the pattern *more X than Y*. This framework is used to identify the point on a semantic scale which most fittingly characterises that which is being described. For example, in sentence 39a, *pessimistic* is compared to *optimistic* and given precedence. Similarly, sentence 39b gives precedence to *right* over *wrong* and sentence 39c to *dead* over *alive*. When a framework such as *more X than Y* is used, antonyms are effectively brought closer together on their given semantic scale. Each context suggests that identifying the correct point on the scale is difficult and that the decision to favour one antonym over the other is marginal. Thus, one cannot be entirely pessimistic over the Third World, but one should be *more pessimistic than optimistic*. Similarly, although Dr Higgs was *a lot more right than wrong*, the clear suggestion is that she was, indeed, wrong to some degree. Sentence 38c is more unusual in that *dead* and *alive* are traditionally regarded as non-gradable antonyms, a label which the comparative nature of this example appears to over-ride. Nevertheless, each of the sentences above can be said to feature a direct comparison between antonyms: a concept is described as being *more X than Y* to give greater precision to its description.

Indirect Comparison

40a Training would be based upon rewarding good behaviour, because behaviourists, Skinner argued, had found that **reward** is more effective than **punishment**.

40b The **new** bills are more colourful than the **old** ones, with designs in green, yellow, blue, orange, red and blue instead of just green and brown.

40c There is also a clear Third World parallel with Dr Kathleen Kiernan's finding that girls who do **badly** at school are four times more likely to become teenage mothers than those who do **well**.

41a But the Labour idea that income tax hits the **rich** harder than the **poor** is also . . . wrong.

41b Strangely, it is easier to **agree** with Mrs Thatcher than to **disagree** with her, such is her domination of the political agenda.

41c For a special occasion, though, a whole tail each is a pleasant surprise, and as a reward for the extra work involved, **small** monk tails are cheaper than **large** ones.

Instead of comparing antonym directly with antonym, the sentences above compare antonyms against a separate, specified scale. This makes their archetypal framework *X is more {adj} than Y* or *X is {adj+er} than Y.* For example, sentence 40a compares *reward* to *punishment*, but does so in terms of their effectiveness. This makes the comparison less direct than those in sentences 39a–c, and this is reflected by the grammatical structure of sentence 40a – the antonymous pair are here pushed further apart and interrupted by an adjective (*effective*). The corresponding adjective of sentence 40b is *colourful,* which provides a scale against which *new bills* can be compared to *old ones*. Sentence 40c is different in that it specifies the exact proportion of teenage mothers *who do badly* at school relative to the number *who do well* (4:1 in favour of the former), but a form of indirect comparison is still taking place.

The majority of Comparative Antonymy sentences discussed so far contain *more* and *than*, which are reliable signals of Comparative Antonymy. However, sentences 41a–c do not contain *more* because the comparative form of the adjective is used (*harder*, *easier*, *cheaper*). This transformation does not affect meaning (Indirect Comparison is still being expressed in these examples), but it does alter the syntactic framework of each context and, therefore, has implications for the antonym retrieval methods to be discussed in Chapter 10.

Preferential Comparison

42a Superficially, her work for Vanity Fair was a departure from her days on Rolling Stone – colour supplants monochrome, people become **active** rather than **passive** adjuncts – but underneath (and this is what Conde Naste was counting on) the message was the same.

42b Wanting to be **happy** rather than **sad**, I accepted – then realised I had nothing to wear.

42c If it has, you will be forced to dig a hole in rock-hard ground and plant the thing, where it will die **slowly** rather than **quickly**.

The sentences above also fail to include *more*, but this is because they adhere to an *X rather than Y* structure, effectively stating a preference for one antonym over the other. One could classify these examples in two distinct ways. First, one could regard the framework *X rather than Y* as being similar to a framework such as *X as opposed to Y* or *X not Y*, arguing that the *rather than Y* part of this structure is simply included to reinforce X by negating its antonym. This position is supported by the 'removability' of *rather than Y* in each of the above sentences. For example, sentence 42a could simply refer to *active adjuncts*; the addition of *rather than passive* may be a strategy to augment *active* in the same way that one might have selected *as opposed to passive* or *not passive*. This suggests that *X rather than Y* contexts may be more at home in the class of Negated Antonymy.[1]

However, this interpretation overlooks other, important aspects of *X rather than Y* and I would argue that the framework is not synonymous with *X as opposed to Y* or *X not Y*. It is similar in that it signals that a distinction has been made between X and Y, but this distinction seems more subtle in the case of the former; *X rather than Y* still reflects some element of comparison, especially when considered in its literal sense. It also features *than*, the most reliable lexical signal of Comparative Antonymy. Arguably, this makes it more analogous with sentences belonging to this class than sentences belonging to the class of Negated Antonymy. The *rather than* part of the sentence is not there to make the opposition more extreme, but to highlight that a choice between antonyms has been necessary.

Equal Comparison

43a However, the **educated** are just as likely to sanction discrimination in the workplace and social life as the **uneducated**.

43b All fat, **unsaturated** no less than **saturated**, is fattening.

43c The Lord Chancellor, who is responsible for legal aid, is introducing a rule that **unmarried** couples living together should be treated for financial assessment purposes in the same way as **married** couples.

The final sub-class of Comparative Antonymy, as exemplified above, features sentences in which antonym is compared with antonym, but no distinction is made between the two in the text. For example, sentence 43a indirectly compares *the educated* to *the uneducated* in terms of their likelihood to sanction discrimination, concluding that the former is *just as likely* as the latter. This is clearly a comparison, even though it is a comparison which yields no

distinction. However, it should not be assumed that *the educated* and *the uneducated* are presented in perfect equilibrium. This sentence is trading on our expectation of inequality. As readers, we expect (or, at least, we are expected to expect) *the educated* to be less likely to sanction discrimination than *the uneducated*. Therefore, this is another example of a 'surprise antonym', following those encountered in Coordinated Antonymy sentences.[2] It would seem that antonyms are often used to defy our preconceptions, regardless of which class they belong to.

Indeed, just as examples of Preferential Comparison nudge towards the class of Negated Antonymy, so examples of Equal Comparison meet criteria for Coordinated Antonymy. For instance, sentence 43a could easily have been presented in a typical coordinated framework ('both the educated and the uneducated sanction discrimination'), and sentences 43b could have claimed that 'all fat, saturated and unsaturated, is fattening'. Such frameworks would have given no explicit syntactic priority to *the educated* and to *unsaturated*, but these terms would still have been implicitly regarded as the 'surprise antonym'. Sentence 43c, in which *unmarried* and *married* are presented in a comparative context which expresses no distinction, is not so blatant in defying our preconceptions, but the fact that a new rule is being introduced does suggest that an inequality has previously existed. Therefore, it can be seen that sentences 43a–c are not unlike those examples of Coordinated Antonymy which are weighted unevenly. However, some form of comparison is still taking place (as evidenced by the phrases *just as likely*, *no less than*, and *in the same way as*) and these examples have duly been assigned to Comparative Antonymy.

To summarise, Comparative Antonymy is the third-largest class of antonymy and is best defined as a phenomenon in which one antonym is gauged against the other, usually to express dissimilitude, occasionally to express similitude. The most common signals of Comparative Antonymy are *more* and (especially) *than*. However, these words are syntactically variable and can fill any of a number of positions in a sentence.

Comparative Antonymy can be broken down into four groups. Direct Comparison refers to the straightforward balancing of one antonym against another. This usually follows the pattern *more X than Y,* a strategy used to signal that the appropriate point on a scale is difficult to identify and is best described in terms of being closer to one end of the antonymous scale than to the other. Indirect Comparison takes a pair of antonymous concepts and compares them along another, distinct scale. Kipling's assertion that 'the female of the species is more deadly than the male' is a well-known example.[3] Preferential Comparison deals primarily with sentences which make use of an *X rather than Y* construction. Such contexts could be interpreted in more than one way, but I have chosen to focus on the fact that, when taken literally at least, a comparison is made between two antonyms and preference is given to X. Finally, Equal Comparison describes a sub-set of sentences which compare two antonymous concepts and conclude that they are equal. Such sentences

have been assigned to Comparative Antonymy because distinction is not always the consequence of comparison. Data suggests that the point of such contexts is often to make explicit that one antonym is the 'surprise antonym'.

Distinguished Antonymy

Definition:	The co-occurrence of an antonymous pair within a framework that alludes to the inherent semantic dissimilarity of those words.
Typical frameworks:	*the difference between X and Y* *separating X and Y* *a gap between X and Y*

Approximately 5.4 per cent of database sentences have been classified in terms of Distinguished Antonymy, making it the fourth-largest new class. Sentences belonging to this class overtly refer to the semantic distinction between antonyms. However, the writer's intention is not only to signal that a distinction is present, but to use this distinction as part of a larger statement. Thus, antonyms act as parameters of a difference or gap. This can be achieved in one of two ways: either the distinction is metalinguistic or the distinction is metaphoric.

Metalinguistic distinction

44a But far from that, Mortimer's father had not given him even a basic moral education, such that today he still doesn't know the difference between **right** and **wrong**, or so he said.
44b But it made the point that the division between **gay** and **straight** is one of many rifts in our society.
44c Scientists admit that the discrepancies between **male** and **female** brains may be less important than education and experience.

45a However, British Petroleum welcomed the increase in the differential between **leaded** and **unleaded** fuel.
45b This blurred distinction between **fact** and **fiction** has long undermined the credibility of ichnology – the study of tracks and traces left behind by extinct animals.
45c Indeed the difference on grain imports between **fast** and **slow** economic growth is greater than the difference between **fast** and **slow** population growth.

The six examples above all refer to the difference between an antonymous pair. This reference is metalinguistic because the writer presupposes our familiarity with 'opposites' in order to make a more general statement. The framework used in these sentences is [*noun phrase*] *between X and Y* where the noun head is *difference* or a synonym thereof: sentence 44a refers to the *difference between right*

and wrong; sentence 44b to the *division between gay and straight*; and sentence 44c to the *discrepancies between male and female brains*. It would appear quite common for writers to allude to the semantic difference between 'opposites' in the mental lexicon and, as such, antonyms are often employed to mark the boundaries of this difference.

Sentences 45a–c exemplify this phenomenon further. The noun head of sentence 45a is *differential*, which allows the difference between *leaded and unleaded fuel* to be discussed. Sentence 45b is slightly irregular in that it refers to *a blurred distinction*, but it is still a *distinction* (blurred or otherwise) which acts as the head of the noun phrase and describes the inherent semantic difference between *fact* and *fiction*. Sentence 45c may initially seem to be a more complex example, but it firmly belongs in the class of Distinguished Antonymy. The example is unusual because it refers to a distinction between the antonyms *fast* and *slow* on two separate occasions: first when they pre-modify *economic growth*; second when they pre-modify *population growth*. These two differences, when applied to grain imports, are presented in comparison to one another, the former being greater than the latter, according to the text. Therefore, the major contrast of this sentence is actually between *economic* and *population*, as these are the variable lexical items amid constant lexical items (*the difference {on grain imports} between fast and slow X growth*). The parallelism present in this sentence is reminiscent of the combination of constant and variable often found in Ancillary Antonymy examples; the difference being that here antonymy is used to cement the 'given' information rather than signal the 'new' (Winter 1982: 110). However, the primary function served by *fast* and *slow* in this sentence is to distinguish between growth speed. It is coincidental that the antonymous pair also contribute towards a greater contextual parallelism in the sentence.

Though all six sentences are metalinguistic in the sense that a distinction between antonyms is overtly referred to, it is important to note that the focus of these examples is always on the difference arising between antonyms, not on the antonyms themselves. For example, sentence 44a is not concerned with telling us that *right* and *wrong* are semantically distinct (this information is treated as 'given' by the text); its aim is to discuss this distinction in terms of how well (or badly) it is understood by a given person. Similarly, the communicative purpose of sentence 44b is not to signal that *gay* and *straight* are different, but to make the point that this difference is one of many rifts. A possible counter-example is sentence 44c, which discusses *discrepancies between male and female brains*. This distinction, considered in isolation, is not necessarily 'given' information, even though it is presented as such. Many readers would not be aware of such discrepancies (although referring to 'discrepancies between male and female' would be more acceptable).

All of the Distinguished Antonymy sentences presented so far have adhered to the framework [*noun phrase*] *between X and Y*. However, the distinction between antonyms need not always be preceded by a noun phrase; the examples below show a verb phrase dissecting the antonymous pair.

46a Mr Craxi's fresh-faced deputy, Claudio Martelli, also dissented, saying that 'one must distinguish between **hard** and **soft** drugs'.
46b The forces must no longer discriminate between **married** and **unmarried** partners, the report says.
46c 'We felt we needed to differentiate between **temporary** and **permanent** diminutions', said John Parry, Hammerson's managing director.

47a The wonderful thing about Grenville's work is that although so many of her underlying concerns – about the victimisation of women, the tragic effects of abuse, the problems produced by separating out our **masculine** and **feminine** instincts – are straightforward and almost didactic, the pull of her work is truly fictional.
47b Unless they become cleaner, safer and more attractive, they will spiral into a destructive decline, with their contracting populations increasingly polarised between **rich** and **poor**.
47c She explains how patients are assessed by the triage nurse (tonight, Linsey), and prioritised into **major** and **minor**.

Sentence 46a is similar to other Distinguished Antonymy examples in that a distinction is made between two antonymous expressions (*hard {drugs}* and *soft drugs*). The only difference is that here, instead of a noun phrase before the preposition *between*, we encounter a verb phrase, namely *must distinguish*. This sentence also confirms that antonymous pairs are presented as co-hyponyms because in this context (and others), *hard drugs* and *soft drugs* act as hyponyms of the superordinate term *drugs*. The rest of the sentence talks about distinguishing between them, but that in itself affirms their similarity. In other words, as in Coordinated Antonymy sentences, antonyms are being decontrasted and treated as co-hyponyms here.

The verb phrase in sentence 46b is *discriminate*, which is used to distinguish between *married {partners}* and *unmarried partners*. In sentence 46c, *differentiate* precedes *temporary {diminutions}* and *permanent diminutions*. Whereas the distinction of earlier sentences was part of a larger noun phrase (indeed, antonymy was little more than post-modification in many examples), here, more focus is placed on the act of discrimination itself. In other words, when antonyms are distinguished by a noun, the opposition is presented as being pre-established, 'given' information; but when antonyms are distinguished by a verb, the opposition is presented as being current, 'new' information.

Sentences 47a–c are similar to sentences 46a–c, but use more unusual verbs to differentiate between antonyms: sentence 47a refers to *separating out* masculine and feminine instincts; sentence 47b tells how people are *polarised* between rich and poor; and sentence 47c describes how patients are *prioritised* into major and minor. These examples are lexically and grammatically different from those Distinguished Antonymy sentences examined earlier, but they remain metalinguistic because they refer to the same linguistic phenomenon –

the semantic difference between antonyms. This difference is also reflected in sentences sub-classified as Metaphoric Distinction, but the examples analysed below all rely on some kind of metaphor to signal dissimilarity.

Metaphoric distinction

48a However, the most visionary parts of his speech concerned the very future of the planet at this 'crucial phase in our history', endangered according to the Soviet leader by growing tensions in the third world, mounting damage to the environment, and an ever-widening gulf between **rich** and **poor** nations.

48b 'The gap between the **successful** and the **unsuccessful** must be narrowed', he said, 'by ensuring that education provides a ladder upon which everyone can find a foothold.'

48c Sir John criticised Mr Sykes for failing to keep 'a clear distance between his **public** and **private** activities' and for 'pressing too hard' for the appointment of a particular consultant.

49a Mozart's Symphony No. 40 and Coronation Mass framed performances of Berg's Violin Concerto and Webern's Second Cantata, reinforcing the feeling of an unbroken tradition and overriding all barriers between **old** and **new**.

49b You'll struggle to find a better delineation of the no man's land between **love** and **hate**, especially one in which both protagonists reek of skunk oil extract.

49c The key is that a homeowner would need to show that a new mortgage was no larger than needed to bridge the gap between the price of the **new** and **old** house.

A variation on metalinguistic distinction occurs when the 'd-word' (*distinction*, *division*, *difference*, etc.) of the sentence is replaced by a metaphor. The examples above are syntactically similar to previous illustrations of Distinguished Antonymy, but here distinctions are signalled metaphorically, not literally.[4] For example, an *ever-widening gulf* is how the difference between *rich {nations}* and *poor nations* is described in sentence 48a. Indeed, spatial metaphors are the most commonly used metaphors in such contexts: sentence 48b refers to a *gap* between the successful and the unsuccessful, and sentence 48c describes the differential between public [activities] and private activities as *a clear distance*.

The metaphor of sentence 49a is also spatial, but refers not to an empty space so much as a physical object, a *barrier* between old and new. Sentence 49b extends the metaphor principle further by describing the semantic space between love and hate as *a no man's land*. These examples show that the metaphors chosen to represent the difference between antonyms are various and often innovative. Finally, sentence 49c initially seems analogous with sentence 48b in that a *gap* between two noun phrases is mentioned. However, closer inspection

reveals that this *gap* is not between the *new {house}* and the *old house*, but rather between the (presumably) high and low prices of these houses respectively.

Therefore, the class of Distinguished Antonymy is home to those sentences which refer to a distinction between a pair of antonymous noun phrases. In the majority of examples, this distinction forms part of a larger noun phrase which is used as either subject or object in the clause. In other words, the antonymous dichotomy (*the difference between X and Y*) is not the primary focus of the sentence. However, in a minority of examples, a verb phrase is used to distinguish between antonyms (*to differentiate between X and Y*). This usually means that the distinction itself will become the primary focus of the sentence. Occasionally, this distinction is not signalled by a 'd-word' such as *difference* or *discrepancy*, but by a metaphor such as *gap* or *gulf* (though not always a 'g-word'!). The effect of this is no different to that of metalinguistic examples because the context is still trading on our familiarity with 'opposites' and incorporating this into a larger statement. The productivity of various frameworks associated with this class of antonymy will be investigated in Chapter 10, but it should be noted that Distinguished Antonymy is heavily marked by the preposition *between*.

Transitional Antonymy

Definition:	The co-occurrence of an antonymous pair within a framework that expresses a movement or change from one location or state to another.
Typical frameworks:	*from X to Y* *turning X into Y* *X gives way to Y*

Of the 3,000 database sentences analysed, exactly 3 per cent were attributed to the class of Transitional Antonymy. Such sentences involve a movement from one antonymous concept to another. This movement is expressed by a limited number of syntactic frameworks, the commonest of which is *from X to Y*. This framework will be analysed first, followed by *turning X to Y* and, finally, *X gives way to Y*. These are prototypical constructions only – many sentences do not adhere to these frameworks exactly, but customise them as required.

from X to Y

50a Her film career similarly has lurched from **success** to **failure**, with enormous periods out of work.

50b The atmosphere of the negotiations was tense, discussion uneven, the mood in both camps swung from **optimism** to **pessimism**.

50c To this day I have no problem crossing the boundary from **straight** to **gay** people, because I have a gay father.

The framework *from X to Y* features in sentences 50a–c. It is used to express a change from one antonymous state to another and, in each example, a metaphor is employed to describe this transition. For instance, in sentence 50a, her film career is said to move *from success to failure* and the verb of transition which signals this movement is *lurched*. The corresponding verb of sentence 50b is *swung*, which describes the movement of the mood in both camps *from optimism to pessimism*. Sentence 50c differs in that its antonymous pair is not clinal and that a physical metaphor (*boundary*) is used to express the nature of the transition *from straight {people} to gay people.* This triplet of Transitional Antonymy sentences adhere strictly to a *from X to Y* framework. However, transitional antonyms do not always occur in such tidy syntactic patterns and the following sentences show how the construction *from X to Y* may vary in text.

51a Inflation is a tax which redistributes wealth to the **sophisticated** from the **unsophisticated**.
51b As the developing countries' debt topped dollars 1,000 bn, interest rates rose, and funds continued to be transferred from the **poor** world to the **rich**, most commodities enjoyed a bonanza.
51c A smaller 'displacer' piston then moves the gas from the **hot** part of the cylinder to the **cold**, causing it to contract.

Semantically, sentences 51a–c are very similar to 50a–c. The only difference is that the second trio of examples follow a more complex structure, the archetypal framework *from X to Y* being supplemented by additional noun phrase lexis: definite articles appear; ellipsis is not always employed; and more extensive modification is used. For example, sentence 51a uses the verb of transition *redistributes* to speak about a movement of wealth to *the sophisticated* from *the unsophisticated*, and is unusual because it describes a movement *to the Y from the X* when one would expect the transition to be expressed more congruently by *from the X to the Y*. Sentence 51b, in referring to *the poor world and the rich*, is unusual because *world* appears after the first antonym. One normally finds ellipsis in this syntactic slot and the noun head instead placed after the second antonym (sentence 50c being an example of this), but this may have rendered *the poor* ambiguous, creating a confusion between poor people in general and the poor world. The noun phrase of sentence 51c is even larger. Here, gas is moved *from the hot part of the cylinder to the cold*. This creates a distance of six words between antonyms, but the transition in question is fundamentally similar to those of sentences 50a–c, even though the chosen framework is more grammatically complex.

turning X to Y

52a Just as the Princess has grown, turning **weakness** to **strength**, so, surely must this country's economy continue to grow.

52b There he was, on the trail of the world's most ruthless terrorist, when slices of his own **fiction** started turning into **fact** before his very eyes.

52c Even **hard** currency has turned **soft**.

In each of the examples above, Transitional Antonymy is expressed by a form of the verb *to turn*: sentence 52a features *turning* as a non-finite verb which signals the transition from *weakness* to *strength*; sentence 52b describes how *fiction started turning into fact*; and sentence 52c uses the present perfective form of the verb to note that *hard currency has turned soft*. This final example is unusual in that it essentially puns on its antonymous pair. Neither *hard* nor *soft* are used in a literal sense here – the *hard* of *hard currency* refers to cash and the idiom *turned soft* suggests that some sort of weakness has been shown. However, the co-occurrence of these two antonyms is far from coincidental. The writer is contrasting *hard* and *soft* in their common, antonymous sense even though both terms are here applied metaphorically.

X gives way to Y

53a It predicts that companies will try to keep themselves leaner and fitter, even when **recession** gives way to **boom**.

53b Economic **optimism** has given way to economic **pessimism** since the great tax-cutting Budget in March.

53c I joined Vogue at just the moment when the **old** guard was giving way to the **new**.

Like all sentences belonging to the class of Transitional Antonymy, the examples above refer to a movement between a pair of antonymous noun phrases. The phrase used to express this movement is *gives way to:* in sentence 53a, *recession* and *boom* are placed in succession as the former *gives way to* the latter; in sentence 53b, the verb phrase takes the perfective aspect as *optimism* and *pessimism* (both pre-modified by *economic*) are presented sequentially; and, finally, *gives way* takes the progressive aspect in sentence 53c as *the old guard* is followed by *the new*. All Transitional Antonymy sentences can be said to feature two antonymous concepts conjoined by a verb phrase reflecting transition; in the examples above, this verb phrase is a form of *gives way to*.

So, in Transitional Antonymy sentences, an antonymous pair is used to express the starting point and finishing point of a change or movement. Often, the metaphor of 'journey' will be used to describe this movement. The most common framework of Transitional Antonymy is *from X to Y*, which describes a change from one state to another, as in the expression 'from rags to riches'. Other frameworks include *turning X to Y* and *X gives way to Y*.

Negated Antonymy

Definition:	The co-occurrence of an antonymous pair within a framework that negates one antonym as a device to augment the other.
Typical frameworks:	*X not Y* *X instead of Y* *X as opposed to Y*

Only 2 per cent of database sentences fall into the class of Negated Antonymy. This is surprising given that Negated Antonymy is arguably the 'purest' form of antonymy, its primary function being to generate a sharper contrast between the two words by making explicit their inherent antonymity. The commonest framework of Negated Antonymy is *X not Y*. This will be examined first, followed by an analysis of non-standard constructions which can be used to achieve a similar effect.

X not Y

54a In my opinion, the public has cause for **pessimism**, not **optimism**, about the Government's plans for a radical reorganisation of arts funding.
54b Well, without the combination of an arms race and a network of treaties designed for **war**, not **peace**, it would not have started.
54c They were backed by Mr Renton who said it was 'inevitable' that businesses would make artistic judgements; they want to be associated with **success**, not **failure**.

55a Democracy means more than the right to pursue one's own self-interest – government must play an **active**, not **passive**, role in addressing the problems of the day.
55b However, the citizen pays for services to work **well**, not **badly**.
55c 'If the aim is to reduce unemployment, it is nonsense to sign up to measures which we all know will add to labour costs and will **discourage**, not **encourage**, employment is this country', Mr Fowler said in a scrutiny debate on the Charter.

The classic framework for Negated Antonymy is *X not Y* and this is used in approximately half of all sentences belonging to this class. Sentences 54a–c show antonymous nouns in X-position and Y-position; sentences 55a–c show adjectives, adverbs and verbs filling these positions. However, regardless of which word class the antonyms belong to, the *not Y* part of the structure is grammatically removable. This does not make it semantically redundant though, and one function of *not Y* is to reinforce the X-ness of X by identifying its semantic scale and negating its antonym.

For example, in sentence 54a, *pessimism* is followed immediately by *not*

optimism. This places greater emphasis on *pessimism*, but it also makes the reader more aware of the alternative. It is as though *pessimism* is the marked choice and the negated antonym (*optimism*) signals the more expected alternative. Similarly, the negation of *peace* in sentence 54b draws attention to *war*, but also signals that treaties are, in fact, normally associated with *peace*. However, it is more difficult to argue that *success* is the marked antonym in sentence 54c because no business would wish to be associated with *failure*. In this example, the function of the second antonym is to emphasise that business not only needs to be associated with *success*, but also needs to avoid being associated with *failure*.

Sentence 55a post-modifies *active* with *not passive*, a strategy which again defies our expectations (or, at least, alerts us to the possibility that governments have been playing a passive role). Sentence 55b uses *not badly* to affirm *well* and to acknowledge (and counteract) any suggestion that the services should work *badly*. And sentence 55c accentuates *discourage* by adding *not encourage*, a move which raises our awareness of the fact that the measures in question were presumably intended to encourage employment, according to Mr Fowler. In the majority of Negated Antonymy sentences, the X-position antonym is the 'surprise antonym'.

As with most classes of antonymy, not all examples adhere precisely to an archetypal framework. The following sentences each make use of *not*, but are more syntactically complex and resist a standard *X not Y* structure.

56a 'What we have to do', Mr Reed says, 'is ensure we shift the balance to the left-hand column, so we get the **good** things out of this, not the **bad**.'

56b Why the champions of A-level cannot get it into their thick heads that the exam guarantees **low** standards, not **high** standards, is beyond me.

56c The case for treating animals better is so intellectually convincing that ours is not a cause to **win**, ours is a cause to **lose**.

The antonymous pair of sentence 56a is part of a larger noun phrase*, the X/Y things out of this*. As such, *good* and *bad* are further apart, syntactically, than antonyms previously examined. Similarly, in sentence 56b, *standards* appears after *low* as well as *high* where one might expect to find ellipsis. Sentence 56c achieves a stylistic effect by negating the clause *ours is a cause to win*, then repeating it verbatim with *lose* in the place of *win*. This is typical of the parallelism associated with spoken political rhetoric,[5] though as the *not* clause appears first in this example, ellipsis is not possible.

An interesting feature of Negated Antonymy sentences (especially those hinging on *not*) is that they tend to reflect spoken rather than written English. Of the nine examples considered so far in this section, two pairs of antonyms (those in sentences 55c and 56a) appear in speech marks. However, all of the other seven examples give the impression, to a greater or lesser degree, that

they are also based on the spoken word. For example, sentence 54a begins with *in my opinion*; sentence 54b begins with the discourse marker *well*; and sentence 54c features a quote from Mr Renton.[6] Similarly, sentence 56b, if not spoken, evidences a strong colloquial style with its reference to *thick heads* and sentence 56c, as mentioned, is reminiscent of political rhetoric. Possible counter-examples are sentences 55a and 55b, but even these contexts represent subjective opinions rather than statements of fact. This suggests that Negated Antonymy is more common in speech and speech-like, persuasive writing than it is in formal writing. But why should this be the case? Perhaps because Negated Antonymy is in one sense redundant;[7] negating the antonym of a word is strictly tautological. The effect of *X not Y* is to create additional rhetorical emphasis, and the immediacy of this emphasis may make it more suited to speech than writing.

Non-standard frameworks

57a Certainly the whole concept of managing by **punishment** instead of **reward** has become part of our culture.
57b If the Germans had met him **alive** instead of **dead**, they would have seen a figure they associated more with Africa than Europe: lean, tattooed armed, wrapped in a cloak of cured skins.
57c Instead of thinking **short** term, it was time to start thinking **long** term.

58a Being **young** and keen, as opposed to being **old** and keen, Stewart wanted to bowl CD out twice and win by an innings.
58b It is a challenge which says, **implicitly** if not **explicitly**, that the Fifth Estate itself is a piece of journalistic mythical self-creation the justification of which is questionable in reality.
58c We are shown a delightful poster of Margaret Thatcher **dressed**, or rather **undressed**, as a whip-wielding madame.

Negated Antonymy need not always be expressed by an archetypal *X not Y* framework. The sentences above achieve the same effect by using alternative constructions. For example, sentence 57a places the nouns *punishment* and *reward* in an *X instead of Y* framework and sentence 57b places the adjectives *alive* and *dead* into the same framework. Sentence 57c is slightly different in that the negated antonym appears first in the clause *instead of thinking short term*. However, all three of these sentences are comparable with those following an *X not Y* pattern because the negated antonym remains removable from its context.

Sentence 58a uses the construction *X as opposed to Y* to much the same ends as previous Negated Antonymy frameworks. Here, *old* is negated to enhance the youngness of *young*. The proximity of *and keen* to each antonym gives stability to the opposition, illustrating further that parallelism and antonymy are common linguistic bedfellows.

Sentences 58b and 58c are more marginal examples of Negated Antonymy. The former sentence follows *implicitly* with *if not explicitly*, a strategy which is difficult to explain. On the one hand, *if not explicitly* is grammatically removable from its context, which makes it analogous with *not explicitly*; on the other hand, the function of *if not explicitly* is very different from the function of *not explicitly*. The key word here is *if*. One could express this difference by stating that *implicitly, not explicitly* is 'emphatically implicit', but that *implicitly, if not explicitly* is synonymous with 'at least implicitly and possibly explicitly'. Sentence 58c uses antonymy with more humorous effects, initially using the verb *dressed*, then replacing it with *undressed*. In this odd example, the antonyms *dressed* and *undressed* function almost synonymously (as either would suffice in this context) and the 'correction' of lexis does not create opposition but, rather, highlights a morphologically unlikely similarity between the two antonyms.

However, in most cases, if a word is post-modified by its negated antonym, the effect is to strengthen the position of that word on its given semantic scale: to make X seem more X-like by rejecting Y. Often, Y will be the unmarked alternative, its inclusion perhaps signalling that it is the more obvious (but instantially dispreferred) option. Negated Antonymy sentences favour the framework *X not Y* (or variations thereof), indicating that *not* is a consistent lexical signal of this class. Other syntactic patterns include *X instead of Y* and *X as opposed to Y*. Finally, the sixty-two database sentences assigned to Negated Antonymy indicate that this function of antonymy may be more common in spoken language than written language.

Extreme Antonymy

Definition:	The co-occurrence of an antonymous pair within a framework that unites the outer-most areas of their given semantic scale.
Typical frameworks:	*the very X and the very Y* *either too X or too Y* *deeply X and deeply Y*

The database features forty sentences which have been classified in terms of Extreme Antonymy, suggesting that the average antonymous pair has a one in seventy-five chance of serving this function in text. The most significant characteristic of Extreme Antonymy sentences is the semantic distance placed between antonyms. Essentially, an adverb pre-modifies each antonym, stretching the contrast further apart on its given scale. The result is a noun phrase which is not exhaustive or inclusive, but which unites two semantic areas (one at each end of the scale) and refers to them rather than to the semantic territory in between. Nine examples follow:

59a It is often considered a safer and gentler form of treatment, especially valuable to the very **young** and the very **old**, being less toxic and having fewer side effects.

59b No-one can afford to go to law except the very **rich** and the very **poor** and it can't possibly get any worse.
59c For thousands of years in Britain, food had to be either very **cold** or very **hot**, but now they are accepting warm salads.

60a The advantages are that the track does not need watering, and can be used when conditions are either too **dry** or too **wet** for racing on turf.
60b You can use it as a planting medium, to improve soil that is too **light** or too **heavy** or as a mulch around ornamental plants.
60c Nothing, it seemed, was too **large** or too **small** for Mr Al-Fayed: opening hours, music, uniforms, design, retailing theory.

61a Freud maintained in Civilization and its Discontents that human beings feel a deep **hate** and a deep **love** for civilization.
61b The Greeks understood so well that there is no pure **masculine** or pure **feminine** in one person; in the poetry of their lives they accepted homosexuality and bisexuality, whose impulses they regarded as just another stream which flowed toward the same great sea – the eternal source of love.
61c Hsu Chu-chuan, the secretary of Hsi Yu, a fishing village, admitted: 'I am not completely **afraid**, and not completely **unafraid**.'

Structurally, the sentences above are reminiscent of those belonging to the class of Coordinated Antonymy. Approximately half of Extreme Antonymy sentences link antonyms with *or* and approximately half link antonyms with *and*, a distribution similar to that of Coordinated Antonymy. Another similarity is *either*, a lexical signal of Coordinated Antonymy which also arises in some Extreme Antonymy contexts. However, Extreme Antonymy is semantically distinct from Coordinated Antonymy. Sentences belonging to the latter class use antonymy to signal inclusiveness or exhaustiveness: often, an entire semantic scale is represented by a pair of antonyms; at other times, the antonymous pair will be attached to a noun phrase to signal that it is an included part of that concept. Extreme antonyms are fundamentally different because their function is to represent the outer reaches of a semantic scale only, not the entirety of that scale. Extreme Antonymy sentences are outnumbered by Coordinated Antonymy sentences at a ratio of 30:1.

The examples above are sorted according to the adverb which gives them their extremity. Sentences 59a–c pre-modify their given antonyms with *very*. For example, the reference in sentence 59a to *the very young and the very old* identifies two areas on the scale of age, but does not encompass the entire scale (as, say, 'everybody, young and old' would). Sentence 59b is slightly different in that it distinguishes between two groups of people: those who can afford to go to the law (*the very rich and the very poor*) and those who cannot (everybody else). As such, *very rich* and *very poor* are instantial co-hyponyms and the

contrast of this sentence is effectively between them and those people who fall nearer the centre of the wealth scale. An interesting aspect of sentence 59c is that a mid-point (*warm salads*) is actually specified by the text. *Very cold* and *very hot* mark either extremity of the 'food temperature' scale, although the major contrast in this context is actually temporal, between *for thousands of years* and *now*.

The next trio of sentences use *too* as their adverb of extremity, which results in an element of negativity: sentences 60a refers to track conditions being *too dry or too wet*; sentence 60b refers to soil being *too light or too heavy*; and sentence 60c refers to nothing being *too large or too small* for Mr Al-Fayed. Once again, in each example, the two extremes of the semantic scale are brought together and contrasted with all that lies in between.

The final trio of sentences use adverbs other than *very* and *too* to signal extremity. The implication of Freud's assertion in sentence 61a is that human beings have no feelings of neutrality towards civilisation, only *a deep hate* and *a deep love*. Sentence 61b is similar to sentence 59c in that two extremes are negated (there is *no pure masculine or pure feminine*) and a central area is identified. The *warm salads* of this example are *homosexuality and bisexuality* (so to speak). Finally, in sentence 61c, the adverb *completely* allows both extremes of the 'fear' scale to be negated.[8] Thus, *deep*, *pure* and *completely* can be seen as lower-frequency alternatives to *very* and *too*.

To conclude, Extreme Antonymy shares many grammatical similarities with Coordinated Antonymy but remains semantically dissimilar (and much less frequent). The key characteristic of Extreme Antonymy sentences is that the two outer areas of a semantic scale are presented in terms of their similarity, often being set up in opposition against the rest of the scale. One could label this the 'Goldilocks Effect' – each extreme is united in its rejection, just as Goldilocks rejected the too-hot porridge and the too-cold porridge; the too-big chair and the too-small chair; and the too-hard bed and the too-soft bed, all in favour of something in between. The most popular framework for this class is *too X or/and too Y*, though *too* is substitutable for *very* or any other adverb which signals extremity.

Idiomatic Antonymy

Definition:	The co-occurrence of an antonymous pair within a framework that would be recognised as a familiar idiom, proverb or cliché.
Examples:	*penny wise and pound foolish* *easy come, easy go* *through thick and thin*

The class of Idiomatic Antonymy accounts for antonyms which feature in an idiomatic expression. The database contains twenty-three such sentences (only 0.8 per cent), but this figure might have been larger had different pairs been

selected. Antonyms not investigated here include *ups*/*downs* and *ins*/*outs*, which hardly ever function in a non-idiomatic fashion. In other words, the frequency of idiomaticity is not stable among individual antonymous pairs and can vary from 0 per cent to 100 per cent. Below are six database examples:

62a The **long** and the **short** of it is that height counts.
62b The critical point is that banks are understandably reluctant to throw **good** money after **bad**.
62c They evidently knew they could teach this **old** dog a few **new** tricks.

63a These abstract pieces seem to glow **hot** and **cold** and seduce by their brilliance.
63b Whoever said the **female** of the species was more deadly than the **male** hadn't met Lord William Whitelaw.
63c Sondheim's aptitude to polarise audiences is distressing to all concerned until everyone gets to know a new piece well enough to **agree** to **disagree**.

Sentence 62a features *long* and *short* in a coordinated context, but so clichéd is this context that we no longer process it in terms of inclusiveness. As sentence 62a is comparable with examples of Coordinated Antonymy, so sentence 62b is comparable with examples of Transitional Antonymy. However, the combination of the framework *throw X money after Y* and the antonyms *good* and *bad* make this example highly idiomatic. Finally, sentence 62c relies on an ancillary pair (*old*/*new*) to signal that *dog* and *tricks* are instantially contrastive, but, once again, such is our exposure to this adage, we do not need to unpack the metaphor to glean its meaning. In other words, although these sentences could be assigned to various larger classes of antonymy, to do so would be naïve; each context is idiomatic and therefore digested as a whole, not as the sum of its individual parts.

Some writers play on our familiarity with idioms to give an unexpected angle to a hackneyed expression and sentence 63a demonstrates this by replacing 'to blow hot and cold' with 'to *glow* hot and cold'. This is because, one assumes, the abstract pieces in question are glowing in some respect. Sentence 63b is idiomatic because the quote from Kipling is so familiar that it has assumed proverbial status, and sentence 63c is idiomatic because the word string *agree to disagree* is also becoming enshrined in the language and something of a cliché. The construction *X to Y*, where X and Y are antonymous verbs seems restricted to *agree*/*disagree* and *love*/*hate*, because most verbs are syntactically unable to occupy X-position.

The majority of sentences assigned to the class of Idiomatic Antonymy feature a conventional idiomatic phrase which includes an antonymous pair. A small number of sentences use word play or puns to bring new life to these old platitudes. The decision to isolate these sentences rather than attribute them to larger classes was taken because the idiomaticity of these expressions overrides the original textual function of the antonymous pair.

Residual database sentences

Having concluded a taxonomy of the six minor categories of antonymy, only 106 database sentences remain unaccounted for. These residual sentences show antonymy functioning in unusual ways. Patterns of usage do emerge, but these patterns are weaker than those identified so far. This section will report on some of the roles that antonyms serve at lower levels of frequency and, where necessary, explain why these sentences have not been assigned to one of the eight larger classes. Seven low-level functions of antonymy will be outlined to provide some insight into the residual component of the database. These functions are presented in order of frequency, but do not collectively account for all 106 sentences. Inevitably, some antonymous usage resists any kind of generalisation.

Conflict

64a For the latter, the Poindexter trial was simply a **good** guy versus **bad** guy quarrel.
64b The survey also shows that the environmental movement has won the debate over **public** versus **private** transport.
64c It's the same battle, the **old** guard versus the **new** guard, as you can see in the USSR: cricket is not immune from the politics of real life.

65a One thing she points out is the intriguing clash of **masculine** and **feminine** qualities Shakespeare arranges at the character's first appearance.
65b Three more uniformed men patrolling the garden fence and three police vehicles parked outside illustrated the conflict between her **old** and **new** lifestyles.
65c So the poem seems to concern itself not just with the slave era or with the poverty of the post-colonial time, but with the construction of something from the clash of the cultures **white** and **black**, **rich** and **poor**, animist and Christian.

Eighteen of the 106 residual database sentences feature the co-occurrence of an antonymous pair within a framework that presents antonym in direct conflict with antonym. The first triplet of sentences above pivot around *versus* and present antonyms antagonistically, as signalled by the corresponding noun: *quarrel*, *debate* and *battle*, respectively, in sentences 64a–c. The second triplet of sentences also create a sense of conflict between antonymous noun phrases but achieve this by using a construction other than *X versus Y*. For example, sentence 65a describes an *intriguing clash* between *masculine and feminine qualities*. This antonymous framework would usually be associated with Coordinated Antonymy (*X and Y n*). However, in this context, the antonyms share no sense of inclusiveness; rather, they oppose one another starkly. Indeed, sentence 65b refers to a *conflict* directly when describing the relationship between *old*

{lifestyles} and new lifestyles. Sentence 65c uses the noun *clash* to signal three separate distinctions. The antonymous pairs *white/black* and *rich/poor* are contrasted, and a difference is also noted between *animist* and *Christian*. However, this example is no different from other sentences. Antonymous noun phrases are presented in a hostile context, this encounter usually being marked by *versus*, but occasionally making use of the framework *n between/of X and Y* where n is *conflict* or a synonym thereof.

Oblique stroke

66a In the visceral imagination there appears to be some mistrust of the **alive/dead** distinction itself, some reluctance to accept that what is dead is henceforth and for ever devoid of life.

66b In his sharply written and strongly performed (by Ian Hughes and Pamela Nomvete above) yuppy morality tale, Below The Belt, Daniel Scott reverses **male/female** stereotypes with some skill; I would have liked to have found out where it went to.

66c The programme note promised a strong **masculine/feminine** contrast followed by final synthesis, yet even Kernis at his dreamiest carried a higher electrical charge.

67a Bell has a **love/hate** relationship with the classic, but does not underrate the prestige associated with winning a race better known than any other on the motor sport calendar, which is one reason why the pre-race organisation barely broke its stride when the sport's governing body kicked Le Mans out of the championship following a row over television rights.

67b Sussex's **new/old** boy Adrian Jones was another man of pace to make an immediate impression of an appropriate nature.

67c The US team feel wronged and are **happily/unhappily** letting their opponents suffer for it.

In 17 of the 3,000 database sentences, the antonymous pair were found to be joined by nothing more than an oblique stroke. These sentences are heterogeneous in nature and are connected more by punctuation than function. Though such examples only occur once per 150 instances of antonymy, it is still interesting to examine why antonyms are used in this way and what role is served by the oblique stroke in contemporary writing.

Each of the examples above use their antonymous pair in a slightly different way. In other words, the oblique stroke performs different roles in different contexts. For example, sentence 66a uses *alive/dead* as pre-modification for *distinction*, a highly metalinguistic use of antonymy. Effectively, the context treats the *alive/dead* distinction as 'given' information. Sentence 66b also use its antonymous pair (*male/female*) as pre-modification, but here the noun head is *stereotypes*. This example is given meaning by the verb *reverses*, which acknowledges the semantic difference between the antonymous pair and makes

the context more grammatical – if the oblique stroke is replaced by an *and*, the sentence remains valid, something which is not true of all examples. Sentence 66c is similar to sentence 66a in that an antonymous pair (*masculine/feminine*), joined by a stroke, precedes a noun head signalling disparity (*contrast*).

The phrase *love/hate relationship* occurs in sentence 67a. More than any other antonymous pair in this framework, this expression has become part of our vocabulary. Within this context, Bell seems to love certain aspects of Le Mans but hate others. This is described as a *love/hate relationship*, a phrase which occupies semantic territory between the extremes of *love* and *hate* and incorporates aspects of both. In some respects, this makes it comparable with sentence 67b. Here, one assumes that Adrian Jones has played for Sussex in the past and recently rejoined the team. Hence, he is either a *new boy* or an *old boy*, depending on one's interpretation. Finally, sentence 67c is equally subjective because the adverbs *happily* and *unhappily* are substitutable for one another in this context (making them co-hyponyms in that respect), depending on one's perspective about the situation. In the eyes of the US team, *happily* applies; in the eyes of their opponents, *unhappily* is more appropriate. Thus, the oblique stroke is almost responsible for some kind of ancillary contrast here.

From such limited data, it is difficult to draw conclusions about this function[9] of antonymy. At times, the oblique stroke is used metalinguistically to refer directly to a conceptual distinction. At other times, both antonyms are applicable to the given context, illustrating that one's perspective may determine which (if either) is the more appropriate. Perhaps the only conclusion one can draw with confidence about this clutch of sentences is that they are very infrequent, occurring only seventeen times in the database.

Association

68a The treaty gives stability to an association between **rich** and **poor** which might otherwise be highly variable.
68b After yesterday's decision to allow South Africa to re-enter world cricket, Rice may prove the only tie between the **old** and **new** eras.
68c The links between **employment** and **unemployment** trends are weak for a variety of reasons.

69a If psychosocial thinking has made us rethink the relationship of **male** and **female** on the basis of mutual support and complementarity, is it not about time that we extended the same thinking to the encounter of faiths as well?
69b John Kirwan will be missed more for the stability he would have brought to this uneasy-looking blend of the **old** and the **new** than for the quality of his play, which the selectors clearly felt was no longer up to standard.
69c In that sense it was counter-productive of Panorama to use TV police themes as shorthand and risk muddling **fact** and **fiction**.

The sentences above could be assigned to the class of Coordinated Antonymy or to the class of Distinguished Antonymy. However, though they meet some criteria for both of these larger classes, the noun phrases which describe the relationship between X and Y prevent antonyms from becoming close enough to be attributed to Coordinated Antonymy and prevent antonyms from being prised apart enough to be attributed to Distinguished Antonymy.

The first triplet of sentences each feature the preposition *between*, a familiar signal of Distinguished Antonymy, where, it may initially seem, these three sentences belong. It is certainly true that they follow closely the framework {*noun phrase*} *between X and Y* where X and Y are antonymous. The only problem here is that the noun phrase is not the word *difference* or a synonym thereof (although one might counter that a distinction of sorts must be presupposed for *between* to be present at all). An *association* is not the same as a distinction, nor is a *tie* and nor are *links*. In many respects, these terms are the antithesis of a distinction, for instead of highlighting the disparity between antonyms, they focus on the connection. Likewise, the *relationship* and the *blend* between antonyms mentioned in sentences 69a and 69b do not serve a differentiating role. Rather, they describe a more complex association, also reflected by the *muddling* of fact and fiction in sentence 69c. Indeed, these examples reflect an association between noun phrases which arguably makes them analogous with Coordinated Antonymy contexts. However, this larger class is not appropriate either because inclusiveness is not being signalled in these sentences ('the relationship of *both* male and female' would sound very awkward in sentence 69a); rather, antonyms are being set up in some kind of correlation which is specific to the given context. Therefore, these sentences (seventeen of which appear in the database) may share grammatical similarities with two larger classes, but they fail to meet necessary semantic criteria so remain in the residual component of the database.

Specification

70a The centre was originally designed to hold 511 men in single cells, but it now houses over 600 **male** remand prisoners and 140 **female** inmates in separate wings.

70b A total of 2,103 drug users, 2,031 **male** and 72 **female**, had contracted Aids by the end of February, David Mellor, the Health Minister told MPs.

70c When the riot began there were 51 **male** and 140 **female** prisoners inside the remand centre.

In sixteen residual sentences, the antonymous pair are quantified. However, sentences 70a–c are different from those Ancillary Antonymy sentences which feature a quantitative B-pair[10] because no element of contrastive power is being generated. Although numbers feature in each of the examples above,[11] these numbers do not appear to be in contrast with one another. For example,

sentence 70a specifies the quantity of *male* and *female* inmates in a prison (*over 600* and *140* respectively). But these figures are not set up in opposition here; they are simply included to provide further information. Similarly, the numbers of *male* and *female* drug users in sentence 70b are not contrastive, and the breakdown of *male* and *female* prisoners in sentence 70c is intended to provide additional detail rather than to express an opposition.

Therefore, antonyms in the sentences above may initially appear to be serving an ancillary role, but, on closer inspection, are not involved in any contrast-generating context. The line between such sentences and those Ancillary Antonymy sentences which feature quantification is sometimes difficult to draw, but the antonymous pair in the above sentences serve a taxonomic role, not an ancillary role.

Simultaneity

71a But that **strength** could also be a **weakness** in that, internationally, he never really accepted that what was good for France was not necessarily equally good for the UK and the US.

71b As one senior Bank of Italy official remarked, 'Mr Amato's **weakness** is his **strength**.'

71c Hyperspace is beautifully written, making **difficult** scientific ideas seem accessible, almost **easy**.

Occasionally, antonyms are directly equated with one another to create an unlikely or ironic parallel. Such contexts do not present antonyms in terms of their similarity (unlike Coordinated Antonymy sentences), but suggest that, in a given context, the dual properties of X and Y may be applicable to the same referent. Only eight examples of this phenomenon were found in the database, three of which are recorded above.

Sentence 71a describes a previously mentioned *strength* as a *weakness*. This apparent contradiction is licensed by perspective – we are told that, internationally, his persuasiveness (for wider context reveals that it is his persuasiveness which is under discussion) is a *weakness*; however, nationally (one assumes), it is a *strength*. In this sense, the antonymous pair of sentence 71a could even be seen as being indirectly ancillary to a second contrast. Sentence 71b features the same antonymous pair and is also difficult to analyse. In effect, *weakness* and *strength* are here operating along different semantic scales. Mr Amato's *weakness*, one assumes, is a personal weakness, but his *strength* is a political strength. As such, the sentence functions on two levels: on the surface, antonym is equated with antonym, but, underneath, those antonyms do not directly contradict one another. Thus, no real paradox is created. Sentence 71c describes a noun phrase (*difficult scientific ideas*) as being almost *easy*. One could argue that Transitional Antonymy is being expressed here, but, once again, the contrast between antonyms is licensed by differing perspectives: the ideas are usually *difficult*, but here seem *easy*.

A similar grammatical structure is adhered to in sentences 71a–c: *X is Y*, where X and Y are antonymous noun phrases and *is* can be replaced by another form of *to be*, or a copula verb such as *seem* (in which case contexts refer to appearance of simultaneity rather than true simultaneity). Such sentences are too rare to justify being classified alone, but it is interesting to note that even functions of antonymy which are very infrequent tend to favour certain frameworks in text.

Unity

72a In most wars, of course, **right** and **wrong** are not nearly so clear-cut, whatever the warring tribes may think.

72b Friends of the Sudan may feel that the issues of **war** and **peace**, of famine and banditry are more pressing, but seemingly for most Sudanese the Sharia remains the central issue that must be resolved before other issues can be tackled.

72c Abuladze's reputation is, however, based on a trilogy of films that deal with fundamental questions of **good** and **evil**, **love** and **hate**, **life** and **death**.

The sentences above initially seem identical to those belonging to Coordinated Antonymy because each antonymous pair is conjoined solely by *and*. Indeed, an element of inclusiveness can also be attributed to these examples, yet this inclusiveness is so familiar that it seems almost over-inclusive. Antonyms have been brought so close together in these contexts that they function as single multi-word units; as one large noun phrase rather than two small noun phrases.

This intuitive reason for not attributing these sentences to the class of Coordinated Antonymy is confirmed by the fact that coordinated antonyms linked by *and* can sustain the word *both* before the *X and Y* phrase. However, this would create an uneasiness in the sentences above; to speak of issues of '*both* right and wrong' or '*both* war and peace', for example, would sound odd. Sentence 72c presents three consecutive units, (*good and evil*, *love and hate* and *life and death*[12]), but the principle remains the same. Abuladze's films do not deal with those six issues individually; rather, they deal with the three concepts represented by each antonymous pair.

The seven database sentences which comprise this residual function of antonymy tend to be signalled by the lexical item which precedes the antonymous pair, usually a term such as *issues*, *questions* or *notions*. This word fills the *n*-slot in the archetypal construction *n of X and Y*. The unity created by this noun, and by our repeated exposure to the antonymous pair, leaves these sentences in a no man's land somewhere between the classes of Coordinated and Idiomatic Antonymy.

Equivalence

73a Dorothy Richardson herself has said that Pilgrimage, her extended autobiographical novel, was 'an attempt to produce a **feminine** equivalent of the current **masculine** realism'.

73b Then there is the possibility that the Hairy Hands story is the **rural** version of the **urban** folk-myth – everyone has heard the one about the deep-fried chicken that was really a battered rat – which plays on people's fears about their environment.

73c To the side, a trio of women swoon or cower with their children, the **feminine** antithesis to this display of steely **masculine** resolve.

Those residual sentences which serve the function of equivalence tend to follow the structure *X n of Y*, where n is *equivalent* or a synonym thereof. Three of the five database examples are recorded above, beginning with sentence 73a, which refers to an X *equivalent* of a noun phrase pre-modified by Y. Sentence 73b is similar but uses *version* to describe the relationship between the antonymous pair. Therefore, a contrast is being set up between antonyms, but this contrast is used only to reflect an underlying similarity. In other words, the similarity of sentence 73b is the *folk-myth*, but the antonymous pair describe the one way in which the *rural {folk-myth}* and the *urban folk-myth* differ.

Sentence 73c is something of a rogue sentence, being the opposite of the previous pair of examples in many respects. Instead of *equivalent* or *version*, this sentence makes use of *antithesis* to describe the relationship between *feminine* and *masculine*. However, it is interesting to note that this example is grammatically identical to sentence 73a, following the pattern *feminine* [*noun head*] *of/to masculine n*, where n is *realism* in sentence 73a and *resolve* in sentence 73c. Of course, sentence 73c cannot technically be described in terms of equivalence (only in terms of a lack thereof), but it is semantically unique in the database and it does share grammatical similarities with valid examples of this function.

Summary of low-level functions of antonymy

The seven residual functions of antonymy outlined above account for 87 of the 106 database sentences which could not be attributed to any of the larger classes. In most cases, the reason that such attribution was not possible may seem trivial. It is true that many of these sentences share similarities with more widespread functions of antonymy, but in each case some important criterion (or criteria) failed to be met. For example, those residual sentences which express conflict (*X versus Y*, etc.) could be seen as a sub-class of either Distinguished Antonymy or Coordinated Antonymy. However, the fact that they can lay claim to both of these classes suggests that they may not truly

belong to either. Distinctions are made between antonyms, but, unlike standard examples of Distinguished Antonymy, this distinction is always in the context of a conflict; and though antonyms are sometimes coordinated, too much of their antonymity remains active for them to be compatible with standard examples of Coordinated Antonymy. Those residual sentences linked by an oblique stroke are diverse in nature and could have been assigned to a number of larger classes. This punctuation mark is multi-functional, but it seems that the act of using it to separate antonyms is itself a rhetorical device. Database sentences which function in terms of association are less complex, and it could be argued that they belong either to the class of Coordinated Antonymy or Distinguished Antonymy. Coordinated Antonymy is also a possible home to those residual sentences which express simultaneity, though they also meet criteria for Ancillary Antonymy. A similar pattern arises with regard to those residual sentences grouped together in terms of unity, rightful ownership of which could be claimed by Coordinated Antonymy or Idiomatic Antonymy. Finally, the class of Ancillary Antonymy could be regarded as home to residual sentences which reflect specification and equivalence, though it is debatable whether the A-pair are truly ancillary (in the case of the former) and whether the B-pair are textually visible (in the case of the latter).

In other words, the main reason why the 106 residual database sentences were kept away from their corresponding larger classes was to allow the criteria developed for those classes of antonymy to remain robust. However, it should be acknowledged that a more cavalier (less pedantic?) approach to classification would allow for the eight larger classes to become larger still. This would give the impression that the textual function of antonymy is even more regular and predictable than already demonstrated.

With regard to the tiny proportion of residual sentences which resisted any kind of classification (nineteen in total; just 0.6 per cent of the database), these examples merely illustrate that corpus linguistics is not an exact science and that antonyms, like all words, are occasionally used in obscure and innovative ways about which generalisations are not easily made.

Major classes vs. minor classes

This chapter has examined what antonyms do in text on the minority of occasions that they are not serving a coordinated or ancillary function. Six minor classes of antonymy have been described and some attention has been given to residual database sentences. Of those six minor classes, four show antonyms functioning in a 'pure' sense, and serving the kind of textual functions that pre-corpus linguists intuitively associated with antonymy. These classes are Comparative Antonymy, Distinguished Antonymy, Transitional Antonymy and Negated Antonymy, all of which could be said to use antonymy in a way that exploits the oppositeness of word pairs to the maximum. Negated Antonymy sentences make the distinction between antonyms as overt as one can imagine by using 'opposites' in a self-conscious, metalinguistic fashion,

effectively spelling out the relationship of antonymy. Distinguished Antonymy sentences also prise 'opposites' apart and capitalise on their semantic dissimilarity. Similarly, Transitional Antonymy and Comparative Antonymy sentences presuppose our awareness of the semantic dichotomy between antonyms; how could a transition or comparison be possible if no difference (or perceived difference) existed?

In other words, these four minor classes use their antonymous pairs in much the way that one might anticipate, exploiting a simple two-way distinction. Yet these four categories collectively account for just 17 per cent of all database sentences. Coordinated Antonymy alone has double this frequency. And yet Coordinated Antonymy does not focus on the distinction between antonyms. Rather, it takes for granted our awareness of this disparity and uses it to signal something else, namely that a given noun phrase is inclusive or exhaustive in some respect.

Like Coordinated Antonymy, Ancillary Antonymy is also twice as common as Comparative, Distinguished, Transitional and Negated Antonymy in total. And like Coordinated Antonymy, Ancillary Antonymy also does more than draw a simple distinction between a pair of antonyms. Contrasts are set up in Ancillary Antonymy sentences, but these contrasts are designed to draw attention to a new, further opposition, usually arising between two expressions which would not otherwise be interpreted in terms of their dissimilarity.

The two minor classes of Idiomatic Antonymy and Extreme Antonymy are different. The former defies generalisation – like most vocabulary, antonyms sometimes crop up as part of a well-known phrase or cliché. However, Extreme Antonymy is essentially a special kind of Coordinated Antonymy. These two classes are syntactically alike, but semantically different. In the opinion of this analyst, distinct classes are necessary, but Coordinated Antonymy could be seen as the parent class of Extreme Antonymy. And, as noted in the previous section, Coordinated Antonymy is also closely related to a number of functions served by residual database sentence. This suggests that the two dominant classes could be even more dominant than this analysis indicates, and that some minor functions of antonymy, though arguably the most intuitively obvious, could be even rarer.

In short, antonymy tends not to appear in a simple, contrastive context; such distinctions may be considered too obvious to draw. Writers prefer to use antonymy as a signal, be it to encompass a scale or to highlight a nearby contrast. Merely stating that X is different from Y is felt to be an inefficient use of language, corpus data would suggest.

7 The endemicity of antonymy

One of the questions which any detailed investigation of antonymy would be remiss to ignore is this: exactly how widespread is antonymy in language? One would intuitively expect antonymy to be highly pervasive because 'opposites' are encountered so extensively in everyday life. However, an accurate assessment of the ubiquity of antonymy is very difficult to establish. First, there is the problem of defining antonymy: the broader the definition one uses, the more widespread the phenomenon will appear. Then there is the even greater problem of counting: in order to arrive at an estimate of the pervasiveness of antonymy within a corpus, one would need to identify every single antonymous pair in use, then retrieve every sentence which features both of those words,[1] then manually edit all of these sentences (which would number over a million in my corpus) to eliminate those in which the word pair do not function antonymously. Only in this way could one arrive at an approximation of the proportion of corpus sentences which feature antonyms, and this approximation would still fail to account for inter-sentential antonymous usage.

An easier way to estimate the prevalence of antonymy in text is to compare the expected co-occurrence rate of antonyms with their observed co-occurrence rate. Justeson and Katz (1991) conducted a similar experiment using the forty antonymous pairs identified by Deese (1964) and the one-million-word Brown corpus. This enabled them to calculate observed/expected (O/E) co-occurrence rates for individual pairs from which an average for all antonyms could be deduced. Justeson and Katz's research has now been duplicated using the new index of fifty-six antonymous pairs and a 280-million-word corpus. This allows statistical comparisons to be made and provides a framework to make judgements about the degree to which antonymy is endemic in English. This chapter will also consider the usefulness of co-occurrence data as criteria for indicating which pairs are 'good opposites'. A handful of such pairs will be identified which, it will be argued, may lay claim to being among the most dyed-in-the-wool antonyms in English.

Antonymy and co-occurrence

Table 7.1 lists all fifty-six antonymous pairs under scrutiny. Each pair is followed by seven columns of figures: columns 1 and 2 record the raw frequency

Table 7.1 Co-occurrence statistics for antonymous pairs

W1/W2	*W1 freq*	*W2 freq*	*Exp*	*Obs*	*O/E*	*W2/O*	*W1/O*
active/passive	11,411	2,033	1.8	172	95.4	11.8	66.3
advantage/disadvantage	21,531	2,483	4.2	69	16.6	36.0	312.0
agree/disagree	18,196	2,472	3.5	153	43.8	16.2	118.9
attack/defend	43,395	9,198	31.0	273	8.8	33.7	159.0
cold/hot	16,466	16,026	20.5	751	36.6	21.3	21.9
correct/incorrect	10,529	1,484	1.2	34	28.0	43.6	309.7
dead/alive	32,214	11,661	29.2	565	19.3	20.6	57.0
deny/confirm	7,514	6,595	3.9	335	87.0	19.7	22.4
difficult/easy	54,244	31,395	132.4	434	3.3	72.3	125.0
directly/indirectly	14,172	1,377	1.5	492	324.3	2.8	28.8
drunk/sober	4,730	1,878	0.7	56	81.1	33.5	84.5
dry/wet	10,978	5,109	4.4	348	79.8	14.7	31.5
encourage/discourage	12,586	1,614	1.6	77	48.8	21.0	163.5
end/begin	145,438	19,682	222.5	740	3.3	26.6	196.5
explicitly/implicitly	1,320	813	0.1	32	383.5	25.4	41.3
fact/fiction	78,900	7,391	45.3	503	11.1	14.7	156.9
fail/succeed	10,963	8,258	7.0	131	18.6	63.0	83.7
fast/slow	22,625	17,374	30.6	350	11.5	49.6	64.6
feminine/masculine	1,191	903	0.1	140	1,674.4	6.5	8.5
good/bad	181,876	47,247	668.1	4,804	7.2	9.8	37.9
guilt/innocence	4,229	3,804	1.3	162	129.5	23.5	26.1
happy/sad	28,217	9,420	20.7	140	6.8	67.3	201.6
hard/soft	68,635	11,960	63.8	526	8.2	22.7	130.5
high/low	93,232	41,088	297.8	2,847	9.6	14.4	32.7
honest/dishonest	6,922	1,084	0.6	28	48.0	38.7	247.2
legal/illegal	40,832	11,208	35.6	302	8.5	37.1	135.2
light/heavy	36,832	22,898	65.6	297	4.5	77.1	124.0
long/short	131,582	52,119	533.2	2,168	4.1	24.0	60.7
love/hate	42,541	6,108	20.2	511	25.3	12.0	83.3
major/minor	45,452	10,624	37.5	432	11.5	24.6	105.2
male/female	16,930	14,883	19.6	2,556	130.5	5.8	6.6
married/unmarried	25,581	1,033	2.1	101	49.2	10.2	253.3
new/old	341,832	113,065	3,004.8	9,426	3.1	12.0	36.3
officially/unofficially	6,025	394	0.2	33	178.8	11.9	182.6
old/young	113,065	83,247	731.8	2,704	3.7	30.8	41.8
optimism/pessimism	5,717	1,163	0.5	91	176.0	12.8	62.8
optimistic/pessimistic	7,123	1,984	1.1	96	87.4	20.7	74.2
permanent/temporary	10,413	7,878	6.4	351	55.0	22.4	29.7
poor/rich	34,054	20,999	55.6	2,027	36.5	10.4	16.8
prove/disprove	20,968	258	0.4	35	83.2	7.4	599.1
public/private	133,056	61,202	633.1	6,741	10.6	9.1	19.7
publicly/privately	8,108	6,406	4.0	282	69.8	22.7	28.8
punishment/reward	6,363	6,152	3.0	38	12.5	161.9	167.4
quickly/slowly	25,129	8,958	17.5	83	4.7	107.9	302.8
recession/boom	22,707	8,678	15.3	334	21.8	26.0	68.0
right/wrong	125,712	42,376	414.2	2,677	6.5	15.8	47.0
rightly/wrongly	4,558	2,681	1.0	182	191.6	14.7	25.0
rural/urban	8,600	7,923	5.3	515	97.2	15.4	16.7
small/large	86,908	69,219	467.7	2,928	6.3	23.6	29.7

Table 7.1 (*continued*)

W1/W2	*W1 freq*	*W2 freq*	*Exp*	*Obs*	*O/E*	*W2/O*	*W1/O*
straight/gay	21,672	9,734	16.4	277	16.9	35.1	78.2
strength/weakness	19,866	5,971	9.2	441	47.8	13.5	45.0
success/failure	47,816	24,438	90.8	971	10.7	25.2	49.2
true/false	35,357	10,245	28.2	227	8.1	45.1	155.8
war/peace	81,293	38,258	241.8	2,586	10.7	14.8	31.4
well/badly	178,431	15,772	218.8	712	3.3	22.2	250.6
win/lose	76,372	27,771	164.9	1,125	6.8	24.7	67.9
Total	**2,662,409**	**955,994**	**8,439.5**	**55,411**			
Average					**6.6**	**17.3**	**48.0**

of each antonym in the entire 280-million-word corpus (higher-frequency antonym (W1) in column 1; lower-frequency antonym (W2) in column 2); column 3 records the number of corpus sentences one would mathematically expect to feature both antonyms if those words co-occurred at random; column 4 records the number of corpus sentences which, in reality, contain both antonyms; column 5 records the observed/expected ratio, which is generated by dividing the figure in column 4 by the figure in column 3; column 6 records the W2/observed ratio, which is generated by dividing the figure in column 2 (i.e. the frequency of the less-common antonym) by the figure in column 4; finally, column 7 records the W1/observed ratio, which is generated by dividing the figure in column 1 (i.e. frequency of the more common antonym) by the figure in column 4. The relevance of each column will now be discussed in turn.

W1 and W2

The first two columns of the Table 7.1 show the raw frequency figures for each antonym. For example, *active* appears 11,411 times in the 280-million-word corpus, and *passive* appears 2,033 times. It may seem surprising that the frequency of *active* is fivefold that of *passive*. However, such discrepancies are not unusual. Indeed, six database pairs are made up of two words, one of which is over ten times more frequent than the other, as recorded in Table 7.2.

The antonymous pair with the most extreme variation is *prove/disprove*. The frequency of *prove* outnumbers the frequency of *disprove* by 81:1. *Unmarried* and *unofficially* can also be regarded as low-frequency derivations of their root woods, occurring twenty-five and fifteen times less often in the corpus than *married* and *officially* respectively. Indeed, the most striking connection between the antonymous pairs in Table 7.2 is morphology. Four of the pairs are morphologically related antonyms and this correlation is strengthened by the fact that the other two morphological pairs in the database (*agree/disagree* and *correct/incorrect*) both show similar, though less extreme, distribution patterns. In each case, the seed word outnumbers its morphological antonym

Table 7.2 Antonymous pairs with most uneven individual frequencies in the corpus

W1/W2	*W1 freq*	*W2 freq*	*W1:W2 ratio*
prove/disprove	20,968	258	81:1
married/unmarried	25,581	1,033	25:1
officially/unofficially	6,025	394	15:1
well/badly	178,431	15,772	11:1
fact/fiction	78,900	7,391	11:1
directly/indirectly	14,172	1,377	10:1

by a wide margin, signalling that antonyms created in this way are much less common in language than the words from which they are derived.

Morphology may account for frequency variation among four of the six pairs recorded in Table 7.2, but why does *well* outnumber *badly*, and *fact* outnumber *fiction* so dramatically? The answer to this question revolves around the number of other contexts in which these words may occur. For example, *well* can function adverbially (and therefore antonymously with *badly*), but it can also be frozen as part of an idiom and is able to adopt other grammatical forms which could not possibly yield antonymy with *badly*, such as adjective, noun and interjection.

74a The **well** of explanation will never run dry, since the question is essentially unanswerable.
74b Early in his career he made what were revolutionary decisions; first, that to learn more about human physiology and disease one should study humans rather than animals, and to study humans one had to be qualified to look after them whether they were ill or **well**.
74c His Soviet Union squad seemed, **well**, asleep; Yugoslavia had scored the first six points of the game and opened a 24–12 lead.

75a The show will include designs in hand-woven silk checks, as **well** as others in elaborate silk jacquards and taffetas.
75b Like an elephant with a broken back, it heaved and twitched and floundered helplessly, **well** and truly trapped.
75c The Home Secretary appoints us and the budget comes from the Treasury, but both Home Secretaries under which we have operated have gone to considerable lengths to stand **well** away from us.

None of the corpus sentences above would allow the substitution of *well* for *badly* to create antonymity. The *well* of sentence 74a is a noun and no more than a homonym of the *well* which contrasts with *badly*. The antonym of *well* in sentence 74b could be *ill* or *unwell* but not *badly*, while sentence 74c shows *well* operating in a colloquial sense as an interjection. Sentence 75a features the common phrase *as well as*, which reflects non-antonymous adverbial usage. It

would not be grammatically incorrect to substitute *badly* here, but the phrase *as well as* is not usually the 'opposite' of the phrase *as badly as*. The *well* of sentence 75b operates as part of an idiom and sentence 75c confirms that, even when functioning adverbially, *well* is by no means guaranteed to be interchangeable with its antonym.

It is for this reason that the frequency of *well* is so much greater than that of *badly* and it is for the same reason that the frequency of *fact* is so much greater than that of *fiction*. In the latter case, corpus data shows that *fact* is usually preceded by either *in* or by a determiner; *fiction* is more limited in its grammatical scope and can only substitute for *fact* in a minority of sentences. This also helps to explain why *married* is more common than *unmarried*. Aside from morphological factors already discussed, *unmarried* is restricted to adjectival usage, but *married* can also function as a verb. Indeed, as one would expect, multi-functional words tend to arise more frequently in the corpus. Thus, *attack* outnumbers *defend* because it can function as a noun as well as a verb and *end* outnumbers *begin* for the same reason.

One final connection which might be made between the six pairs of words with most uneven individual frequencies is that a disproportionate number of them are adverbial. However, this may be coincidental. Although *badly*/*well*, *directly*/*indirectly* and *officially*/*unofficially* all appear in Table 7.2, other adverbial pairs in the database share more equal individual frequencies (for instance, *explicitly*/*implicitly*, *privately*/*publicly* and *rightly*/*wrongly*).

Indeed, it is interesting to pause for a moment to consider those word pairs at the opposite end of the scale which occur at similar rates in the corpus. For example, the raw frequency of *cold* (16,466) and *hot* (16,026) are remarkably alike, and other antonymous pairs whose individual frequencies differ by less than 10 per cent include *punishment*/*reward*, *rural*/*urban* and *guilt*/*innocence*. One could argue that the similarity of frequency between *cold* and *hot* gives this pair greater claim to the status of being 'good opposites' than the extreme dissimilarity of frequency between, say, *disprove* and *prove* because the potential antonymity of *disprove* is usually dormant when *prove* is used, but the potential antonymity of *cold* is usually active when *hot* is used. However, as demonstrated earlier, the fact that a pair of antonyms occur at uneven rates in the corpus does not necessarily detract from their antonymity. Rather, this could be indicative of homonymy (e.g. *light* meaning 'bright' as well as 'weightless'), polysemy (e.g. *straight* meaning 'linear' or 'heterosexual'), or grammatical fluidity (e.g. *right* operating as an adjective, noun or adverb). This does not prevent such pairs from being 'good opposites'.

Expected co-occurrence

Using the raw frequency data in columns 1 and 2 of Table 7.1, it is possible to calculate the number of sentences which could be expected to feature both members of each antonymous pair if words were somehow able to co-occur in a mathematically arbitrary fashion. For example, were *alive* and *dead* in no

semantic relationship, one would expect them to appear together in about twenty-nine corpus sentences. This figure is calculated by multiplying the chance of *alive* appearing in a given sentence by the chance of *dead* appearing in a given sentence. The corpus contains 12,862,340 sentences, so the chance of *alive* appearing in a random sentence is 11,661 divided by 12,862,340, which equals just over 0.09 per cent, about once per 1,100 sentences. The chance of *dead* appearing in a random sentence is slightly greater: 32,214 divided by 12,862,340, which equals just over 0.25 per cent, about once per 400 sentences. If we multiple these two likelihoods together, we find that *alive* and *dead* can be expected to occur together about once per 440,000 sentences (1,100 × 400). As there are 12,862,340 sentences in the corpus, this means that about 29 of these sentences (12,862,340/440,000) can be hypothetically expected to feature both words.[2]

Returning to the fifty-six word pairs examined, the range of expected co-occurrence figures is enormous. Because *new* and *old* each appear in over 100,000 corpus sentences, their expected co-occurrence rate exceeds 3,000. At the other end of the scale, because *explicitly* and *implicitly* each appear in under 1,500 corpus sentences, their expected co-occurrence rate is less than 0.1. Calculating an expected co-occurrence rate for each antonymous pair is useful because it allows a comparison to be made between the number of corpus sentences one would expect to feature both antonyms and the number of corpus sentences which actually do feature both antonyms.

Observed co-occurrence

Column 4 of Table 7.1 shows the observed co-occurrence rates for each antonymous pair, i.e. the number of corpus sentences in which both antonyms actually appear. Once again, a wide range of figures are recorded here. At the bottom end of the scale, a dozen antonymous pairs appear in fewer than 100 corpus sentences, the lowest being *dishonest* and *honest*, which co-occur in just twenty-eight corpus sentences. At the top end of the scale, a further twelve antonymous pairs appear in more than 1,000 sentences, the highest being *new* and *old* which co-occur in 9,426 sentences. The top ten word pairs of the sample, in terms of co-occurrence frequency, are listed in Table 7.3

This table provides an insight into which antonymous pairs are most favoured in contemporary journalistic text. The most striking aspect of the above list is that it seems very adjective-heavy. Nine of the ten pairs are adjectival and the sole exception (*peace*/*war*) is ranked only eighth. The most frequently occurring verbal pair (*lose*/*win*) would occupy twelfth spot on the list and the most frequently occurring adverbial pair (*badly*/*well*) would be in sixteenth place. To some degree, the high levels of co-occurrence recorded by adjectives justifies the attention traditionally paid to antonymous members of this word class by semantic theorists (e.g. Palmer (1972) and Jackson (1988)).

If frequency of co-occurrence is seen as a criterion of 'good opposites', then *new* and *old* can rightfully claim to be the king (and queen?) of antonyms. By

Table 7.3 Top ten rates of intra-sentential corpus co-occurrence among antonymous pairs

1	new/old	9,426
2	private/public	6,741
3	bad/good	4,804
4	large/small	2,928
5	high/low	2,847
6	old/young	2,704
7	right/wrong	2,677
8	peace/war	2,586
9	female/male	2,556
10	long/short	2,168

some margin, they are the most commonly co-occurring antonymous pair examined here, and it is therefore likely that they are the most commonly co-occurring antonymous pair in English. The only pretenders to the crown of *new*/*old* would be a pair of prepositions such as *against*/*for* or *from*/*to*. The former pair co-occur in over 50,000 corpus sentences; the latter co-occur in over half a million. The problem is that neither *against*/*for* nor *from*/*to* function antonymously in the majority of those sentences. More often than not, co-occurrence will be 'accidental' because *for* and *to* (and, to a lesser extent, *against* and *from*) are multi-functional items with extremely high frequency. One cannot assume that such pairs are functioning antonymously just because they co-occur in the same sentence.

The most important aspect of the observed co-occurrence column is that, without exception, the figure recorded for each antonymous pair is greater than the expected co-occurrence for that pair. In other words, all of the fifty-six word pairs analysed co-occur more often than chance would allow.

Observed/Expected co-occurrence

The figures in column 5 of Table 7.1 are calculated by dividing the observed co-occurrence for each pair by their expected co-occurrence. This answers the question: how many times more do these antonyms co-occur than would be expected by chance? For example, *advantage* and *disadvantage* co-occur in 69 corpus sentences, yet were only expected to co-occur in 4.2 corpus sentences. This means that their rate of co-occurrence is 16.4 times greater (69/4.2) than chance would allow.

The range of values generated by the observed/expected calculation is broad, the smallest being 3.1 (*new*/*old*), the largest being 1,674.4 (*feminine*/*masculine*). The enormity of the latter figure can be partially explained by the corpus being too small to adequately generate expected co-occurrence statistics for words which themselves are relatively rare. As the frequency of *feminine* is just 1,191 and the frequency of *masculine* just 903, the number of sentences in the corpus which would be mathematically expected to feature both words is

only 0.0836. Therefore, the corpus would need to be twelve times larger (around 3.3 billion words) before a single sentence could be expected to feature this opposition.

However, the corpus actually contains not one, but 140 sentences which feature both *feminine* and *masculine*, generating an O/E figure of 1674.4. Furthermore, *feminine/masculine* is not the only low-frequency word pair to generate a large O/E ratio: *explicitly/implicitly* also has an expected co-occurrence below 0.1, yet records an observed co-occurrence of 32, resulting in an O/E figure of 383.5; *officially/unofficially* has an expected co-occurrence of 0.2, yet records an observed co-occurrence of 33, resulting in an O/E figure of 178.8.

It is interesting to glance at the extreme figures generated by low-frequency pairs, but, in many cases, the individual frequencies of the words involved are not large enough to make the O/E figure statistically meaningful. Or, rather, the corpus, large as it is by current standards, is not vast enough to support co-occurrence analyses of such infrequent items. Therefore, Table 7.4 lists those word pairs which succeed in recording a strong O/E figure despite being high-frequency items. The table shows the top ten O/E scores for word pairs with an expected co-occurrence rate of five[3] sentences or more.

As Table 7.4 shows, *male/female* generates by far the highest O/E score for any high-frequency antonymous pair. One would expect two words with the raw frequency of *female* and *male* to co-occur in about twenty corpus sentences; however, the pair actually co-occur in over 2500 corpus sentences. Other high-frequency word pairs which co-occur in the same sentence much more than would be mathematically expected are *rural/urban*, *permanent/temporary* and *strength/weakness*, though it should be noted that these O/E scores are derived from lower expected co-occurrence rates than *male/female*.

Although a high O/E ratio could not be seen in isolation as a valid criterion for antonymy, one would be surprised if a pair of 'good opposites' did not score well in this test. Hence, it is reassuring to note that all fifty-six of the antonymous pairs examined record an O/E score of at least 3.0. In other words,

Table 7.4 Top ten observed/expected co-occurrence rates among high-frequency antonymous pairs

		Obs	*Exp*	*O/E*
1	male/female	2,556	19.6	**130.5**
2	rural/urban	515	5.3	**97.2**
3	permanent/temporary	351	6.4	**55.0**
4	strength/weakness	441	9.2	**47.8**
5	cold/hot	751	20.5	**36.6**
6	poor/rich	2,027	55.6	**36.5**
7	love/hate	511	20.2	**25.3**
8	recession/boom	334	15.3	**21.8**
9	fail/succeed	131	7.0	**18.6**
10	straight/gay	277	16.5	**16.9**

every pair of antonyms in the database co-occur in corpus sentences at least three times more often than expected by chance.

The lowest O/E score is recorded by *new/old* (3.1), followed closely by *badly/well*, *begin/end* and *difficult/easy* (all 3.3). However, it would be premature to jump to any conclusions about these pairs being 'poor opposites'. All four antonyms have uneven individual frequencies, the most extreme case being *well*, which arises on 178,431 occasions in the corpus, compared to *badly*, which records only 15,772 hits. This suggests that these words are unfaithful antonyms: that one partner strays into contexts such that antonym substitution is syntactically or semantically unlikely. In other words, because some antonyms are polysemous (or even polygamous, to continue the metaphor), their expected co-occurrence rate could be distorted. And this, in turn, could render their O/E figure artificially low.

Hence, an inverse correlation emerges between frequency and O/E score. High-frequency antonyms tend to record lower O/E figures because they have more scope to be used in ways which leave their potential for antonymy untapped. For instance, it is very difficult to envisage a context in which *explicitly* and *implicitly* are not substitutable for one another, yet words such as *well*, *end* and *easy* will often function in ways such that their status as antonyms is irrelevant. Thus, one could conclude that the more core an item is in one's vocabulary, the less faithful it is likely to be to its antonym.[4]

However, the co-occurrence rate for all fifty-six word pairs remains statistically significant. Using the chi-square method, all pairs record a score of at least 15, the lowest being *quickly/slowly*, which scores 15.6. Any chi-square score over 10.83 is 99.9 per cent significant at one degree of freedom (Oakes 1998: 266). Therefore, high rates of co-occurrence among these antonymous pairs cannot be attributed to chance.

Word 2/observed

The sixth column of figures is calculated by dividing the number of sentences which feature the less frequent antonym (as recorded in column 2 of Table 7.1) by the number of sentences which feature both antonyms (as recorded in column 4 of Table 7.1). This tells us what proportion of W2 sentences also feature W1. So, the W2/O figure for *attack/defend* is 33.7 – the number of corpus sentences which feature *defend* (9,198) divided by the number of corpus sentences in which *defend* co-occurs with *attack* (273). Therefore, approximately one sentence in thirty-three which features *defend* will also feature *attack*. With *bad/good*, the W2/O figure is 9.8, so approximately one in every ten sentences which features *bad* will also feature *good*. The smaller the W2/O figure, the stronger the likelihood that the lower-frequency item will co-occur with its antonym. The lowest ten figures are recorded in Table 7.5.

As this table shows, if a sentence features *indirectly*, the chance of it also featuring *directly* is approximately one in three. This makes *directly* an excellent collocate of *indirectly*. Three other pairs in this list are also morphologically

Table 7.5 Top ten W2/O scores for antonymous pairs

		W2/O
1	directly/indirectly	2.8
2	male/female	5.8
3	feminine/masculine	6.5
4	prove/disprove	7.4
5	public/private	9.1
6	good/bad	9.8
7	married/unmarried	10.2
8	poor/rich	10.4
9	active/passive	11.8
10	officially/unofficially	11.9

related (*disprove/prove*, *married/unmarried*, *officially/unofficially*) and it would seem that antonyms derived in this way tend to cling to their root word more closely in text than lexical antonyms. However, not all of the word pairs in Table 7.5 are morphologically related. Data shows that more than one in six sentences which feature *female* will also feature *male*; and that *feminine/masculine* co-occur at a similar rate. One can only speculate about why gender-based adjectives appear in the same sentence with such regularity. It could have something to do with *female/male* being a non-gradable pair, but other non-gradable pairs (*alive/dead*, *true/false*, *legal/illegal*, etc.) record much higher W2/O scores. Besides, *feminine* and *masculine* are gradable. Indeed, if an adequate explanation can be found to account for the high co-occurrence rates of *female/male* and *feminine/masculine*, it must surely revolve around the semantics of these terms. This is confirmed by the fact that other gender terms also generate low W2/O scores. For example, to take a selection of gender-based oppositions at random, *father/mother* generate a W2/O score of 8.4; *daughter/son* score 8.9; *man/woman* 11.0, and *brother/sister* 11.2. These figures all reflect a co-occurrence rate which is stronger than the average among antonyms, though an exception to this norm is *boy/girl* which scores a less marked 20.2. This indicates that gender-based terms tend to keep closer company with their antonyms in text than non-gender-based terms.

At the other end of the scale, *punishment/reward* scores 161.9, the weakest of any antonymous pair. Perhaps because *reward* acts as a verb as well as a noun in text, it would seem that these words are less drawn to one another than other pairs. This could weaken their status as 'good opposites'. However, even though only one in every 162 *reward* sentences also features *punishment*, it should be remembered that less than one in 2,000 *reward* sentences would feature *punishment* if these words co-occurred at random.

Word 1/observed

The seventh and final column of figures is calculated by dividing the number of sentences which feature the more frequent antonym (as recorded in column

Table 7.6 Top ten W1/O scores for antonymous pairs

		W1/Obs
1	male/female	6.6
2	feminine/masculine	8.5
3	rural/urban	16.7
4	poor/rich	16.8
5	public/private	19.7
6	cold/hot	21.9
7	deny/confirm	22.4
8	rightly/wrongly	25.0
9	guilt/innocence	26.1
10	directly/indirectly	28.8

1 of Table 7.1) by the number of sentences which feature both antonyms (as recorded in column 4 of Table 7.1). This indicates what proportion of W1 sentences also feature W2. The lower the figure in this column, the stronger the tendency for the higher-frequency word to co-occur with its antonym. And the lowest figure recorded by any antonym is that of *male*, a word which occurs on 16,930 occasions in the corpus, about 2,000 more times than *female*. These antonyms co-occur in 2,556 sentences which means that one in every 6.6 sentences which features *male* also features *female*.

At the other end of the scale, the highest figure in this column is recorded for *prove*. Only one sentence per 599 which features *prove* also features *disprove*. This confirms that the pair are unevenly matched antonyms: *prove* is only used in terms of its antonymy with *disprove* in a fraction of contexts; the rest of the time, this antonymity remains dormant. As the next two highest-scoring antonymous pairs are *advantage/disadvantage* and *correct/incorrect*, morphology can be cited as the major contributory factor in this high W1/O score. The ten lowest W1/O ratios are listed in Table 7.6.

It is interesting to compare those pairs which score well in Table 7.6. (W1/O) with those that score well in Table 7.5 (W2/O). Some correlation emerges, but this pattern is not as strong as one might expect. For example, *directly/indirectly* tops Table 7.5 but occupies tenth place in Table 7.6. This is because *indirectly* is very faithful to *directly*, but *directly* is less faithful to *indirectly*. Conversely, *rural/urban* ranks much higher[5] in Table 7.6 than Table 7.5 because these antonyms have similar individual frequencies. Pairs that score well on both indices include *female/male*, *feminine/masculine*, *private/public* and *poor/rich* and this may indicate that these antonyms are among the most deeply ingrained in language.

The average antonymous pair

Using the figures calculated for each individual antonymous pair, an average O/E figure, an average W2/O figure, and an average W1/O figure are calcul-

able. These averages can be used as a basis for more general observations about the phenomenon of antonymy.

Average O/E

The sum of all the figures in column 3 of Table 7.1 is 8,441.8. In other words, one would expect the fifty-six antonymous pairs investigated to co-occur with one another in a total of 8,442 corpus sentences. The actual number of co-occurrence sentences is 55,411. This means that, on average, the fifty-six word pairs under scrutiny co-occur 6.6 times more often than would be expected by chance.

When Justeson and Katz carried out the same experiment on the Deese antonyms, they found that the average antonymous pair co-occurred in the same sentence 8.6 times more often than would be expected by chance (1991: 142). These two figures, calculated independently, are close enough to prove that antonyms do co-occur in text at a relatively high rate. However, the difference between my figure of 6.6 and Justeson and Katz's figure of 8.6 requires further investigation.

One disparity between our methodologies is that Justeson and Katz limit their research to words which function adjectivally, whereas my database includes non-adjectival antonymous pairs. Therefore, one explanation could be that antonymous adjectives record a higher co-occurrence rate than antonymous non-adjectives. This hypothesis was tested by calculating co-occurrence rates for antonyms according to their word class.

As Table 7.7 shows, adjectives do not co-occur any more often than the average for all fifty-six word pairs. Antonyms which are adjectival co-occur 6.1 times more than would be expected by chance. The difference between this figure and the average for all antonyms (6.6) is minor. Verbs and adverbs co-occur slightly more than average, but the most striking difference arises when nouns are examined. The nine antonymous pairs of nouns in the database co-occur, on average, 12.6 times more than expected. In other words, an antonymous noun is twice as likely to co-occur with its partner than an antonymous adjective.

Therefore, this analysis does not account for the difference between average O/E scores (calculated to be 8.6 by Justeson and Katz and calculated to be 6.6

Table 7.7 Observed/expected co-occurrence rates for antonymous pairs by word class

	Observed co-occurrence	*Expected co-occurrence*	*O/E*
Adjectives	45,020	7,327.9	6.1
Nouns	5,195	411.4	12.6
Verbs	3,380	457.1	7.4
Adverbs	1,816	243.1	7.5
Total	**55,411**	**8,439.5**	**6.6**

here). The explanation seems not to lie in Justeson and Katz's adjectives-only policy, but rather in their choice of corpus. The Brown corpus is tagged, so Justeson and Katz were able to restrict their search to adjectival usage of their chosen antonyms. The newspaper corpus on which the new statistics are based is untagged. This creates the problem of 'wastage' – multi-functional words being retrieved as part of, say, an adjectival antonymous pair when one or both words are actually functioning as nouns. The level of 'wastage' in my corpus is impossible to gauge with accuracy, but the expected number of co-occurrences for most word pairs would rise if the raw frequency of each word was adjusted downwards to account for sentences in which antonyms do not belong to the same word class. In turn, this would allow the O/E figure to rise towards Justeson and Katz's estimate of 8.6.

In other words, although I have calculated antonyms to co-occur within the same sentence 6.6 times more often than expected, this may be distorted by multi-functional antonyms and the true figure may be nearer to 8.6. However, in defence of my estimate, it should be noted that the statistics presented here are derived from a much larger corpus than that used by Justeson and Katz. Perhaps, an acceptable compromise would be to state that the average antonymous pair will co-occur intra-sententially about seven or eight times more often than would be expected by chance.

Average W2/O

In addition to calculating an O/E figure, Justeson and Katz also compute a W2/O figure of fourteen. This allows them to state that 'co-occurrences took place on average once per fourteen sentences containing the less frequent member of an antonym pair' (1991: 141). The W2/O figure is calculated by dividing the frequency of the less common antonym by the number of sentences in which both antonyms appear. Therefore, according to Justeson and Katz, antonym co-occurrence takes place once per every fourteen opportunities in text.

By totalling the frequency of the less common antonym of each word pair (955,994) and dividing by the total number of co-occurrence sentences (55,411), I was able to calculate this figure to be 17.3. In other words, according to my data, the average antonymous pair will co-occur about once every seventeen opportunities in text. Again, the difference between my figure and that generated by Justeson and Katz is relatively small and can be largely accounted for by the fact that my corpus is untagged.

Average W1/O

Justeson and Katz did not calculate an average W1/O figure. In one sense, this figure is not as fundamental as W2/O because it does not reflect co-occurrence opportunities taken. However, a brief glance at the differences between the word pairs of Table 7.5 and those of Table 7.6 demonstrates that antonyms which score well on one criterion do not necessarily score well on the other.

The total of column 1 of Table 7.1 is 2,662,409. The average W1/O score is calculated by dividing this figure by the sum of column 4, which is 55,411. The result is 48.0. This means that lower-frequency antonyms co-occur with their partners once per forty-eight opportunities, on average.

'Good opposites'

This chapter has examined the raw frequency of antonyms in a large corpus, generating co-occurrence statistics accordingly. The notion of 'good opposites' is largely subjective,[6] but it may be possible to justify making a handful of initial requirements of antonymy, based on co-occurrence criteria. Below are four ways in which the statistical evidence presented here could be used as a first step towards identifying those antonymous pairs which can best be described as 'good opposites':

- Table 7.3 lists the ten antonymous pairs which co-occur most frequently in the database. Carter notes that 'the less core a word is, the more difficult it is to find an antonym for it' (1987: 36) and any attempt to identify 'good opposites' must acknowledge that a high rate of observed co-occurrence is a key indicator of antonymity.
- Table 7.4 lists those pairs which achieve a high observed/expected ratio. All antonymous pairs examined co-occur in text at least three times more often than chance would allow, and one could reason that the higher the O/E ratio, the more powerful the antonymity of that pair is likely to be.
- Table 7.5 lists those pairs which achieve a low W2/O figure. This represents the number of opportunities for co-occurrences per actual co-occurrence. The lower this score, the greater the chance of the antonym with lower frequency appearing in a sentence alongside its partner. Given that 'good opposites' could be reasonably expected to remain faithful to one another in text, a low W2/O score may be indicative of a high level of antonymity.
- Table 7.6 lists those pairs which achieve a low W1/O figure. This records the number of times in which the antonym with higher frequency co-occurs with its partner, relative to its raw frequency. Essentially, it measures how faithful the more core antonym is to its less core partner. Once again, a low W1/O score may be indicative of a high level of antonymity.

The four criteria above are based solely on raw frequency and co-occurrence frequency and, as such, hardly provide the ideal mechanism for gauging which antonymous pairs are most fundamental in modern English. However, if one applies these four criteria of 'good opposites' to all fifty-six antonymous pairs, only six record a better-than-average score for each test (i.e. they co-occur in over 990 corpus sentences; their observed/expected ratio exceeds 6.6; their W2/O score is below 17.3; and their W1/O score is below 48.0):

- bad/good
- female/male
- high/low
- peace/war
- poor/rich
- private/public.

Based only on crude co-occurrence criteria, one could argue that these six word pairs are the most 'hardcore' antonyms in language, or, more specifically, that they are the most 'hardcore' antonyms in newspaper text (for these pairs are just the kind of everyday 'opposites' that one would intuitively expect to co-occur most forcefully in journalistic corpora). Five of the six pairs are adjectival, only one is non-gradable, and none are morphologically related. Of course, whether it is methodologically sound to identify 'good opposites' exclusively on the basis of a handful of co-occurrence criteria is debatable. However, it is not unreasonable to assume that all individual antonymous pairs themselves operate along a scale with core, high-frequency, high-fidelity pairs at one end, and less core, lower-frequency, lower-fidelity pairs at the other. The criteria suggested here to identify 'good opposites' are no more than a first step towards the goal of distinguishing between these two extremes of the antonymity scale.

So is antonymy endemic?

This chapter has presented a statistical analysis of antonym co-occurrence in text. All fifty-six word pairs selected for study have been investigated in terms of raw frequency, expected co-occurrence and observed co-occurrence. These figures have been used to calculate the observed/expected ratios and the proportion of actual co-occurrences to possible co-occurrences for each antonymous pair. Based on this data, four criteria for 'good opposites' have been established, and six of the fifty-six word pairs were found to record better-than-average scores for each. Arguably, this gives them some claim to be among the most antonymous word pairs in English.

Unfortunately, this analysis does not yet justify its chapter heading. Co-occurrence statistics prove that antonymy is a ubiquitous feature of language, but the endemicity of antonymy, if antonymy is indeed endemic, remains unproven. To address this, we must begin with the figure of 55,411 (Table 7.1, column 4 total), the aggregate number of corpus sentences which feature both members of one of the fifty-six antonymous pairs selected for study. A further 10,326 sentences can be added to this aggregate to account for the total number of *un*-word sentences in the corpus (i.e. sentences in which antonymy arises because a word co-occurs with the *un*- version of itself). This gives a new total of 65,737: the total number of corpus sentences which definitely feature both members of an antonymous pair. This approximates to one sentence per 195.

However, the true proportion of antonymous sentences is likely to be much greater. The database, though relatively large, incorporates no more than a sample of all antonymous pairs in language. Plurals, past tenses, progressive aspects, comparatives and superlatives are all excluded from the database, even though they too can express antonymy. Also conspicuous by their absence are reciprocal verbs, antonymous prepositions and antonymous pronouns. Furthermore, the index of fifty-six pairs is hardly exhaustive and many other 'opposites' could have been uncontroversially included. Therefore, the true proportion of corpus sentences which feature both members of an antonymous pair could be much greater, perhaps as high as one in fifty. Although this is an unsubstantiated estimate, the demonstrable fact that antonyms co-occur in at least one sentence in 195 is sufficient to evidence some degree of endemicity in text.

8 Antonym sequence

> . . . to have and to hold from this day forward, for **better** for **worse**, for **richer** for **poorer**, in **sickness** and in **health**, to love and to cherish, till death us do part.
>
> (Solemnization of Matrimony, *1662 Book of Common Prayer*)

This chapter seeks to answer the following question: why are marriage vows not taken 'for worse for better, for poorer for richer, in health and in sickness'? In other words, why do antonyms favour a particular sequence in text?

Given that antonyms are equal in many respects (effectively being co-hyponyms and showing 'minimum contrast' (Clark 1970: 275)), one would not necessarily expect pairs to have marked and unmarked sequences in text. There is no obvious reason why, say, *good* should appear before *bad* in a disproportionate number of sentences, nor is there any reason why *bad* should appear before *good* in a disproportionate number of sentences. However, if we examine the 117 database sentences which feature these two words, we find that *good* precedes *bad* 100 times, and *bad* precedes *good* just seventeen times; 85 per cent of sentences which feature both antonyms mention *good* first.

Furthermore, *good*/*bad* is not the only word pair to evidence such a bias: *win* precedes *lose* in 84 per cent of sentences; *rich* precedes *poor* in 86 per cent of sentences; and *male* precedes *female* in 83 per cent of sentences. Indeed, of the fifty-six antonymous pairs investigated, only fourteen do not show any marked preference[1] towards either word. The majority of pairs favour a given sequence over the reverse of that sequence; this chapter will speculate why.

Sequence statistics

This section will identify which antonymous pairs tend to favour a particular sequence in text and how marked that tendency is. Table 8.1 lists all fifty-six pairs in order of the extent to which they adhere to their normal sequence in text: the antonym which is first-mentioned in more database sentences is recorded in column 1; the antonym which is first-mentioned in fewer database sentences is recorded in column 2; the total number of database sentences for each pair is recorded in column 3; the number of those sentences which follow

normal sequence is recorded in column 4; this information is then expressed proportionally in column 5 (i.e. the percentage of database sentences in which normal sequence is observed); and this information is expressed as a binomial score in column 6 (the closer to 1.00 the score, the stronger the correlation and the lower the chance of the distribution being random).

Three of the fifty-six word pairs examined achieve a 100 per cent score: *correct* precedes *incorrect*, *prove* precedes *disprove*, and *rightly* precedes *wrongly* in every database sentence. To some extent, these statistics need to be adjusted to account for the relatively low sample sizes (eighteen, fourteen and forty-four contexts respectively), but all three pairs record a binomial score of 1.000 and one can safely infer that these antonyms show a very marked tendency to favour their normal sequence in text.

Three of the other fifty-three pairs under scrutiny occur in the same sequence within every database sentence but one. Thus, a strong bias is shown by *directly* to precede *indirectly*, *confirm* to precede *deny*, and *officially* to precede *unofficially*. Table 8.1 also shows that the pairs *agree*/*disagree* and *married*/*unmarried* appear in that sequence in 90 per cent of contexts and a further nine pairs occur in the same sequence at a rate of 80 per cent or over. Significantly, these nine pairs include antonyms which are heavily represented in the database. For example, *rich*/*poor* and *good*/*bad* both feature in over 100 sentences yet remain strongly inclined towards mentioning *rich* and *good* first respectively. Exactly half of all antonymous pairs achieve a binomial score of 1.000, indicating that the sequence which they favour is statistically significant to a very high degree.

Further down Table 8.1, a bias of between 60 per cent and 70 per cent is less marked, and a bias of between 50 per cent and 60 per cent could not be seen as marked in any sense. For instance, *easy* precedes *difficult* in 55.6 per cent of database sentences. However, if just two of those database sentences which feature *easy* before *difficult* were replaced by two sentences which feature *difficult* before *easy*, then *difficult* would show a bias of 51.9 per cent. Statistical insignificance is signalled by a binomial score below 0.95 and those fourteen pairs which fall below this threshold cannot be regarded as having (yet) developed a normal sequence and a reverse sequence in text.

However, to return to the other extreme of the scale, Table 8.1 indicates that some antonyms could be regarded as irreversible binomials, pairs of words which favour a given sequence so strongly that to reverse this sequence is highly marked. Examples of irreversible binomials include 'well and truly', 'do or die' and 'nip and tuck'. Following Birdsong (1995), McCarthy (1998) discusses iconicity in the word order of binomials, concluding that binomials, 'because of their idiomaticity, are best taught and learnt as unanalysable wholes' (1998: 148) rather than as the sum of their individual parts. If some antonymous pairs are irreversible binomials, and some irreversible binomials are idiomatic, the database classification undertaken in earlier chapters may require some revision. Sentences which report on a refusal to *confirm or deny* a proposition, for example, have been assigned to the class of Coordinated Antonymy, but McCarthy's comments suggests they may be better assigned

Table 8.1 Sequence statistics for antonymous pairs in the database

A1	*A2*	*Total database sentences*	*Normal sequence (raw freq)*	*Normal sequence (%)*	*Binomial score*
correct	incorrect	18	18	**100.0**	1.000
prove	disprove	14	14	**100.0**	1.000
rightly	wrongly	44	44	**100.0**	1.000
directly	indirectly	79	78	**98.7**	1.000
confirm	deny	34	33	**97.1**	1.000
officially	unofficially	25	24	**96.0**	1.000
married	unmarried	31	28	**90.3**	1.000
agree	disagree	49	44	**89.8**	1.000
true	false	62	55	**88.7**	1.000
legal	illegal	31	27	**87.1**	1.000
right	wrong	60	52	**86.7**	1.000
rich	poor	102	88	**86.3**	1.000
good	bad	117	100	**85.5**	1.000
win	lose	58	49	**84.5**	1.000
boom	recession	24	20	**83.3**	1.000
honest	dishonest	12	10	**83.3**	0.997
male	female	87	72	**82.8**	1.000
well	badly	53	42	**79.2**	1.000
temporary	permanent	28	22	**78.6**	1.000
hot	cold	59	46	**78.0**	1.000
optimism	pessimism	21	16	**76.2**	0.996
success	failure	88	67	**76.1**	1.000
advantage	disadvantage	36	27	**75.0**	0.999
begin	end	51	38	**74.5**	1.000
optimistic	pessimistic	47	35	**74.5**	1.000
love	hate	104	77	**74.0**	1.000
publicly	privately	47	34	**72.3**	0.999
old	new	254	182	**72.0**	1.000
quickly	slowly	28	20	**71.4**	0.994
young	old	69	49	**71.0**	1.000
wet	dry	31	22	**71.0**	0.995
public	private	134	95	**70.9**	1.000
succeed	fail	63	45	**71.4**	1.000
masculine	feminine	68	47	**69.1**	1.000
fact	fiction	36	24	**66.7**	0.986
innocence	guilt	44	29	**65.9**	0.989
high	low	32	21	**65.6**	0.975
happy	sad	45	29	**64.4**	0.982
large	small	50	32	**64.0**	0.984
gay	straight	33	21	**63.6**	0.960
attack	defend	30	19	**63.3**	0.951
reward	punishment	19	12	**63.2**	0.916
active	passive	96	60	**62.5**	0.995
drunk	sober	18	11	**61.1**	0.881
long	short	36	22	**61.1**	0.934
encourage	discourage	28	17	**60.7**	0.908
war	peace	15	9	**60.0**	0.849

Table 8.1 (*continued*)

A1	*A2*	*Total database sentences*	*Normal sequence (raw freq)*	*Normal sequence (%)*	*Binomial score*
minor	major	27	16	**59.3**	0.876
fast	slow	28	16	**57.1**	0.828
light	heavy	77	44	**57.1**	0.914
hard	soft	32	18	**56.3**	0.811
easy	difficult	27	15	**55.6**	0.779
weakness	strength	35	19	**54.3**	0.750
urban	rural	24	13	**54.2**	0.729
dead	alive	54	29	**53.7**	0.752
implicitly	explicitly	30	16	**53.3**	0.708

to the class of Idiomatic Antonymy. The issue of idiomaticity will be discussed later in the chapter, but the first question that will be addressed is: what factors affect the sequence in which antonyms appear?

Sequence rules

Normal sequence among antonymous pairs is not arbitrarily established. The vast majority of pairs observe a sequence which is determined by one or more relevant factors. These factors, listed below in order of influence, form the basis of sequence 'rules' which are obeyed by most antonymous pairs with surprising regularity.

Morphology

The most dominant single factor affecting antonym sequence is morphological derivation. For proof of this, consider the nine antonymous pairs in the database which feature morphologically related antonyms and their tendency to place their root word before their morphological antonym. Table 8.2 shows that all of these nine morphological antonyms show a strong tendency to place their root word before their derivation. The pair with the least pronounced bias are *advantage* and *disadvantage*, which still favour that sequence in text, but not to the same degree as the other morphological pairs – in one quarter of database sentences, normal sequence is reversed and *disadvantage* precedes *advantage*. All other morphological pairs show a bias of 83 per cent or more, with seven achieving a binomial score of 1.000. Final proof of the extent to which morphology affects antonym sequence can be found in Table 8.1: of the ten pairs which record the strongest bias, seven are morphological antonyms.

As one would expect, this pattern is also reflected in the *un*-words portion of the database. Of the 156 sentences randomly selected from the corpus which feature a word co-occurring with the *un*-version of itself, 130 present the root

Table 8.2 Sequence statistics for morphological antonyms in the database

		Normal sequence (%)	*Binomial score*
correct	incorrect	100.0	1.000
prove	disprove	100.0	1.000
directly	indirectly	98.7	1.000
officially	unofficially	96.0	1.000
married	unmarried	90.3	1.000
agree	disagree	89.8	1.000
legal	illegal	87.1	1.000
honest	dishonest	83.3	0.997
advantage	disadvantage	75.0	0.999

word first and the morphological antonym second. This equates to about 84 per cent, which is compatible with the scores for individual pairs in Table 8.2.

In total, the nine word pairs in Table 8.2 feature in 295 of the 3,000 database sentences. Of these 295 sentences, the root word precedes the morphological antonym in 270. Only 8.5 per cent of examples reverse this sequence. If the 156 *un*-word sentences are added, we find that 400 out of 451 database contexts (88.7 per cent) feature their root word before their morphological antonym. This suggests that morphology is a powerful criterion of antonym sequence in text, though it does not explain why morphologically related antonyms adhere so closely to this sequence. To address this issue, other factors must be examined, and, as one would expect, overlaps sometimes arise between these factors. Indeed, morphology may be a symptom of antonym sequence more than a cause, as subsequent sections will show.

Positivity

If one member of an antonymous pair has more positive connotations than the other, it will usually display a tendency to precede its partner in text. Lyons notes that 'the positive opposite tends to precede the negative when opposites are co-ordinated' (1977: 276) and I would add that this observation extends to all intra-sentential antonymous usage, whether coordinated or not.

For example, *good* is a more positive word than *bad*. Therefore, *good* precedes *bad* in 85.5 per cent of database sentences. The corresponding adverbial pair display a similar behavioural pattern, with *well* preceding *badly* in 79.2 per cent of database sentences. For the same reason, *optimism* and *optimistic* appear before *pessimism* and *pessimistic* in 76.2 per cent and 74.5 per cent of database sentences respectively; *success* and *succeed* are both first-mentioned at a rate of over 70 per cent; and *right* and *rightly* precede *wrong* and *wrongly* in 86.7 per cent and 100 per cent of contexts respectively.

A clear pattern can be identified here – antonyms with positive associations are being given priority over antonyms with negative associations. Furthermore, the more positive the associations, the more marked the pattern tends to

be. For instance, other words which occur before their antonymous partner in a high percentage of sentences, thereby achieving a binomial score of 1.000, include *win* (84.5 per cent), *boom* (83.3 per cent), *true* (88.7 per cent) and *rich* (86.3 per cent). These words carry strongly positive connotations, especially if one compares them with their negative partners (*lose*, *recession*, *false* and *poor*). Words which carry less positive associations (but which, none the less, remain positive) include *active* (because *passive* usually suggests inactivity or weakness) and *innocence* (which can carry connotations of naïvety, but remains much less negative than *guilt*, especially in the legal sense). These words are both positive, but their positivity is perhaps less clear-cut than earlier examples. Because of this, although they precede their antonyms in a majority of sentences, that majority is less marked (62.5 per cent and 65.9 per cent respectively).

Other antonyms which tend to occur in a sequence influenced primarily by positivity are *confirm*/*deny*, *happy*/*sad* and *love*/*hate*. More marginal cases include *fast*/*slow* and the corresponding adverbial pair *quickly*/*slowly*.[2] High speed is not always thought of positively (it often has associations with danger), but this is arguably outweighed by the negativity of *slow* and *slowly* (which can be linked to inertia, sluggishness or even unintelligence).

Similar problems are raised by *attack* and *defend*. These words follow that sequence in 63.3 per cent of sentences retrieved, suggesting, perhaps, that *attack* is the more positive antonym. Though the act of attacking is not usually seen in a good light, it could be interpreted as the more active antonym, with *defend* possibly implying inferiority or subordination. A complementary explanation is that these words are often used metaphorically by journalists, in which case *attack* need not carry violent connotations at all – if the thing being attacked is itself considered wrong, that criticism could be seen in a positive light. This may help to explain why *attack* appears first in a majority of contexts.

The sequence of most antonymous pairs is influenced by positivity when one antonym can be identified as being the more positive.[3] However, some counter-examples are unavoidable and four pairs seem to buck the trend of sentences mentioning the more positive antonym first: *dead* precedes *alive* in a majority of sentences; *weakness* precedes *strength*; *war* precedes *peace*; and *drunk* precedes *sober*. To a greater or lesser degree, each of these pairs are occurring with their negative partner before their positive partner.[4] However, the binomial score for all four of these pairs is statistically insignificant and, in general, it would appear that positivity is a key factor in determining which antonym should appear first in the sentence.

Indeed, one could go further and argue that the criterion of positivity subsumes the criterion of morphology. If we re-examine the nine morphological pairs listed in Table 8.2, we can see that the root word is the more positive word in almost every case: *advantage*, *correct*, *honest*, etc. are positive terms; *disadvantage*, *incorrect*, *dishonest*, etc. are negative terms. The possible counter-example is *married*/*unmarried*, which would require a greater degree of subjective judgement before a positive/negative distinction could be made.

So, why has the criterion of morphology been presented separately from the criterion of positivity? For two reasons, the first of which is intimated by the problem of *married/unmarried*. Morphologically derived antonyms are usually negative, but they are not always negative. Consider the antonyms *selfish* and *unselfish*. The negative term here is the root word; the morphological antonym carries more positive connotations. Similarly, *uncensored*, *uncontaminated* and *unbiased* are all frequently presented as positive terms, despite their prefix. Very few corpus sentences feature these words and their corresponding antonyms, but the tendency among those which do is to present the (negative) root word before the (positive) morphological antonym in a majority of sentences.[5] Therefore, the criterion of morphology over-rides the criterion of positivity.

This relates to the second reason for keeping the two criteria distinct, namely that morphology is a more influential factor than positivity. The nine morphological antonyms occur in their normal sequence (root word first; morphological antonym second) in 91.5 per cent of sentences. But positive words precede negative words in only 77.9 per cent of database sentences (the latter percentage is based on all word pairs where a positive and negative term can be identified, including morphological antonyms). This suggests that the criterion of morphology is more powerful than the criterion of positivity and explains why it has been given attention separately.

Magnitude

Morphology and positivity are the two most important factors in determining antonym sequence, but the concept of size also seems to be important in the case of a handful of word pairs. For example, *large* occurs before *small* in 64.0 per cent of database sentences. One might link this with the criterion of positivity (bigger is usually better), but *small* could hardly be seen as a negative term in the same way as *failure* or *hate*. Therefore, the expression of magnitude itself appears to be a valid criterion of antonym sequence.

A second example is the pair *long* and *short*, which follow that sequence in 61.1 per cent of sentences sampled. A third, though perhaps more marginal example is the pair *high* and *low*, which follow that sequence in 65.6 per cent of database sentences. *High* does not reflect magnitude as directly as *large* and *long*, but it could still be seen as the 'greater' antonym when compared with *low*.

A possible counter-example is the pair *heavy/light*. Given the favoured sequence of the pairs above, one might expect *heavy* to precede *light*, as it would usually express the greater magnitude. However, Table 8.1 shows a slight bias to the contrary. *Light* appears before *heavy* in 57.1 per cent of database sentences. Though this bias is not statistically significant, one might account for it by saying that the criterion of positivity is more influential than the criterion of magnitude. Among the things characterised as being *light* or *heavy* in the seventy-seven database sentences are *traffic*, *drinking* and *rain*. In each instance, *light* would be regarded as the preferable, positive alternative, despite being lesser in terms of magnitude.

Chronology

If one antonym is prone to precede the other in the real world, this sequence will be reflected in the syntax of an average sentence. For example, *begin* precedes *end* in three-quarters of all sentences sampled. The obvious reason for this is that the beginning of an event always precedes its end; to reverse this real-world sequence in language would be marked.

This theory is interesting if applied to the pairs *new*/*old* and *old*/*young*. In 72.0 per cent of *new*/*old* sentences, *old* is the first-mentioned antonym. However, *old* is the first-mentioned antonym in only 29.0 per cent of *old*/*young* sentences. This supports the view that, when deciding upon antonym sequence, we are aware of the polysemy of *old* – when we refer to 'not new' *old*, we are likely to record it as the first antonym; but when we refer to 'not young' *old*, we are likely to record it as the second antonym.

The logic behind the sequence of these pairs is perhaps analogous with the logic behind the favoured sequence of *begin* and *end*. Chronologically, *old* precedes *new* because, for example, *old* houses were built before *new* houses. However, chronologically, *young* precedes *old*. Not literally perhaps (because *old* people were born before *young* people), but rather in the sense that one is twenty years old before one is forty years old. Therefore, in text, *old* tends to precede *new*, but *young* tends to precede *old*. This awareness of the temporal aspect of antonymy could help to explain why *attack* tends to precede *defend*, although it would hardly account for the (admittedly less significant) tendency of *dead* to precede *alive*. One might also apply this criterion to *temporary*/*permanent*, arguing that *temporary* precedes *permanent* in 78.6 per cent of sentences sampled because things are sometimes *temporary* before they become *permanent*.

Gender

One of the most striking preferences among antonymous pairs is that of *male* and *female*, with the former preceding the latter in 82.8 per cent of database sentences. One could explain this with reference to the morphological relation[6] between the two words, but the 'prefix' *fe-* is hardly commonplace and *female* is not a low-frequency derivation of *male*, as the relationship between *unmarried* and *married* could be described. The positivity criterion should not apply because it would be sexist to identify *male* as being positive and *female* as being negative[7] (or vice versa), so the only remaining option is to recognise gender itself as a criterion of antonym sequence. This is supported by an examination of the normal sequence of *feminine*/*masculine*. Once again, the bias is not trivial – 69.1 per cent of database sentences feature *masculine* before *feminine*. This seems to confirm the tendency for things *male* to be given priority over things *female* in text, as does the propensity of *he* to precede *she* (even when we use these words in a politically correct fashion to avoid sexist presumption), and, to a lesser extent, *him* to precede *her*, and *his* to precede *hers*.[8]

Phonology

The tendency of *male* to precede *female* in text is so strong that gender alone may not be sufficient to account for it. Given that this sequence is especially noticeable in coordinated contexts such as *male and female* or *male or female*, it is possible that phonological factors are at work here. If we say *male and female*, the phonetic repetition of *male* is interrupted by two syllables; however, if we say *female and male*, the phonetic repetition of *male* is interrupted by only one syllable. Subconsciously, we may wish to keep identical syllables away from one another in speech and, perhaps, text. This would provide another explanation as to why root words precede their morphological antonyms in text.

To investigate the hypothesis that word length influences antonym sequence, those antonymous pairs which differ in syllables from one another (*boom/recession*, *easy/difficult*, etc.) were separated from those which do not (*confirm/deny*, *explicitly/implicitly*, etc.). Twenty-one pairs were found to contain antonyms of different length. Of those, seventeen tend to mention the shorter word first and only four tend to mention the longer word first. Admittedly, nine of those seventeen pairs are morphological antonyms, but even among lexical pairs, twice as many showed a bias towards placing their low-syllable word before their high-syllable word than showed a bias towards the contrary. This may help explain why the following pairs tend to appear with their shorter antonym first: *male/female*, *boom/recession*, *well/badly*, *fact/fiction*, *reward/punishment*, *drunk/sober*, *easy/difficult* and *dead/alive*. Only four of the fifty-six word pairs sampled act as counter-examples: *begin/end*, *succeed/fail*, *innocence/guilt* and *happy/sad*. It could be that other factors (such as positivity or chronology) outweigh phonological factors in these examples.

Idiomaticity

Some word pairs seem to favour a certain sequence, not because of abstract semantic criteria, but because they have developed semi-idiomatic status, perhaps as a result of a certain coinage point in their history. The best example of this is the antonymous pair *war* and *peace*, which tend to follow that sequence despite *peace* being the more positive term. One explanation could be Tolstoy's novel *War And Peace* (1869), which may have set a precedent for those antonyms. Conscious of this in the back of our minds, it is possible that we remain disinclined to reverse this sequence. A second feasible example of idiomaticity affecting sequencing is provided by *alive/dead,* which occurs with the negative term (*dead*) first in a majority of sentences. To a small extent, this could be influenced by the *Wanted Dead Or Alive* cliché of certain films, usually Westerns, and, one assumes, authentic posters of fugitives from American history.

Though the tendencies of *war* to precede *peace*, and *dead* to precede *alive* are not statistically significant (60.0 per cent and 53.7 per cent of database sentences respectively), the difference between this and other overtly

positive/negative pairs is significant (*good* precedes *bad* in 85.5 per cent of sentences, *win* precedes *lose* in 84.5 per cent, etc.). It would be interesting to test the idiomaticity hypothesis by considering the sequence of *peace*/*war* and *alive*/*dead* in pre-twentieth century corpora.

Frequency and markedness

Two criteria which affect the sequence of antonymous pairs to a lesser degree are frequency and markedness. However, as some antonymous pairs seem resistant to these factors, they are presented tentatively and should be regarded as marginal criteria only. Whereas morphology, for example, seems to dictate antonym sequence among all relevant pairs, frequency and markedness are much less exhaustive in their influence.

With regard to frequency, one might expect the more common antonym to 'lead the way' in the sentence because it is the more familiar. However, this hypothesis cannot be conclusively affirmed. In some cases, a strong correlation arises between frequency and antonym sequence: for example, the frequency of *win* is more than double that of *lose*, and *win* precedes *lose* in 84.5 per cent of database sentences; similarly, the frequency of *quickly* is more than double that of *slowly*, and *quickly* precedes *slowly* in 71.4 per cent of database sentences. However, whilst frequency is possibly an influencing factor in determining sequence for these pairs, other antonyms indicate that this correlation should be treated with caution. Consider the pair *new*/*old*: of the 254 database sentences which feature these words, 72.0 per cent feature *old* before *new*. Yet *new* is the higher-frequency word in text, occurring at treble the rate of *old*. Similarly, *rich* appears before *poor* in 86.3 per cent of database sentences, but *poor* is the higher-frequency antonym, recording almost twice as many corpus hits as *rich*.

These counter-examples diminish the case for frequency to be regarded as a criterion of antonym sequence. In total, thirty-six of the fifty-six word pairs examined show a bias towards placing their higher-frequency item first in sentences.[9] However, the other twenty pairs do not, which means that frequency, if an influencing factor at all, belongs towards the bottom of any ranked list of significant criteria.

With regard to markedness, one might expect the unmarked antonym, where one can be identified, to be the first-mentioned antonym in the sentence. McCarthy notes that irreversible binomials can be seen as 'moving from unmarked to marked term in antonymous pairs' (1998: 148), citing *high and low* and *good and bad* as examples. Evidence from the database suggests that markedness[10] is a relevant factor, but that this may be a symptom rather than a cause of the sequence in which antonyms appear. This is because the unmarked antonym is usually the more positive antonymy (*true* is unmarked; *false* is marked) and tends not to be morphologically derived (*legal* is unmarked; *illegal* is marked). In other words, overlaps arise between criteria which determine markedness and criteria which determine sequencing. Moreover, not all antonymous pairs incline towards mentioning their unmarked term

first. Of the sixty-nine *young/old* sentences in the database, the marked antonym, *young*, appears before the unmarked antonym, *old*, in forty-nine (71.0 per cent). Though this pair may be an exception to the rule, the normal sequence of *young/old* demonstrates that markedness does not always determine antonym sequence.

Therefore, although frequency and markedness may affect the sequence favoured by some antonymous pairs, I regard these factors as being less influential than criteria previously discussed. Higher-frequency antonyms show some inclination to precede their lower-frequency partners, and unmarked antonyms likewise tend to appear before marked antonyms. However, problematic counter-examples challenge both of these generalisations and it may be the case that frequency and markedness, rather than actively determining antonym sequence, are themselves subject to the influence of other sequencing criteria.

Problem pairs

Collectively, the above criteria account for the bias displayed by almost all of the word pairs examined. However, a handful of antonyms occur in a textual sequence which (so far) defies explanation. Most of these pairs can be discounted because the bias they show is too insignificant to worry about, especially given the relatively small number of sentences sampled. For example, of the twenty-four *rural/urban* sentences in the database, eleven mention *rural* first and thirteen mention *urban* first. To speculate about this 'bias' is unnecessary; all that has been illustrated is that writers show no strong sequential preference when using these words. Similarly, *implicitly/explicitly*, *weakness/strength*, *soft/hard*, *light/heavy* and *minor/major* all resist conforming to a statistically significant sequence in text.

More relevant is that *hot* precedes *cold* in 78.0 per cent of database sentences and that *wet* precedes *dry* in 71.0 per cent of database sentences. These patterns are difficult to account for in terms of positivity (*hot weather* is usually preferable to *cold weather*, but *wet weather* is rarely preferable to *dry weather*), nor does morphology play any role in their textual sequence. Both pairs of words feature the same number of syllables and magnitude is not applicable. It would seem that these pairs have adopted a sequence which cannot be explained with reference to any of the usual criteria.[11]

Perhaps the least explicable sequence of all antonymous pairs sampled belongs to *private/public* and *privately/publicly*. In 70.9 per cent of sentences examined, *public* precedes *private* and in 72.3 per cent of sentences examined *publicly* precedes *privately*. Therefore, this sequence has become fairly established in text, though, once again, it is difficult to know why (or when) this might have happened.[12] Maybe things public are seen as more important (and therefore more newsworthy) than things private, in broadsheet journalism at least. It would be interesting to discover whether a corpus of tabloid news would be more inclined towards private matters.

The only other pair which favour a sequence that cannot be accounted for by

the factors suggested above is *gay*/*straight*. The tendency of *gay* to precede *straight* (63.6 per cent) is only just statistically significant, but it may be relevant that this is the most recently coined antonymous pair of the fifty-six under scrutiny. The sequence of many antonymous pairs is likely to have become established in centuries long past; and once a sequence is established in language, it is unlikely to be reversed. As *gay* and *straight* are relatively modern antonyms, it is possible that their sequence has been determined by criteria about which it is too early to generalise.

Why are sequence rules sometimes ignored?

Most antonymous pairs in the database favour one sequence over another, but only three of these fifty-six pairs adhere to their normal sequence in every database sentence. And even those pairs that chalk up a 100 per cent score do so only as part of a relatively small sample. In other words, many sentences contain antonymous pairs which break the accepted rules, as they have been outlined above, and reverse normal sequence. Some of those sentences will now be examined with a view to assessing why writers occasionally choose to flout convention.

good *and* bad *in reverse*

At the beginning of this chapter, *good* and *bad* were used to exemplify how most antonymous pairs give syntactic priority to one antonym over the other. Of the 117 database sentences which feature both *good* and *bad*, 100 feature the words appearing in that sequence. But what about those seventeen sentences in which *good* and *bad* occur in their reverse, less conventional sequence?

Of those seventeen sentences, no less than fourteen belong to the class of Ancillary Antonymy. This strongly suggests that the sequence in which antonyms appear is affected by the textual function which they are serving. Three typical examples of reverse sequence Ancillary Antonymy sentences are presented below (76a–c), followed by the three database sentences in which *bad* precedes *good*, but do not form an ancillary pair (77a–c). The first of these sentences belongs to the class of Transitional Antonym; the latter pair to Coordinated Antonymy.

76a Luckily, all these subjects which are very **bad** for *master classes*, are very **good** for *after-dinner speaking*.

76b The Conservative Party may be **bad** at *many things*, but it has been superbly **good** at *winning general elections*.

76c All **bad** medicine for *Britain* and therefore **good** medicine for *the IRA*.

77a 'No amount of fixing', Alan Jay Lerner liked to say, 'can turn a **bad** musical into a **good** one.'

77b It's important to keep hold of the whole gamut of art – we cover the **bad** art as well as the **good**.

77c Speaking generally, I think an English player would not act like that but foreigners are bringing **bad** things as well as **good** things to the game here.

The first triplet of sentences show *bad* and *good* helping to signal a further contrast within each sentence (*master classes* and *after-dinner speaking* in sentence 76a; *many things* and *winning general elections* in sentence 76b; *Britain* and *the IRA* in sentence 76c). The fact that *bad* precedes *good* in these sentences is abnormal, but this abnormality is diminished by the syntactic distance of antonyms from one another. The closer together antonyms are in text, the more likely they are to follow their normal sequence. This is illustrated by the following statistic: among *good*/*bad* sentences attributed to Coordinated Antonymy, only 4.3 per cent occur in reverse sequence; however, among *good*/*bad* sentences attributed to Ancillary Antonymy, 25.5 per cent occur in reverse sequence. The difference between these two proportions is striking, but the latter percentage is still not as large as one might expect. Even among Ancillary Antonymy sentences, sequencing rules are still obeyed because only one-quarter of contexts mention *bad* before *good*. Thus, at some level, our intuitions about word sequence are still operating, even when the antonymous pair are relatively distant from one another in the sentence.

Sentence 77a refers to a transition, in this case *from a bad musical to a good one*. Like Ancillary Antonymy examples, this framework allows antonyms to be syntactically distant from one another. In turn, this diminishes their need to conform to the normal sequence. Conversely, coordinated constructions are least likely to feature their antonyms in reverse sequence. Indeed, of the forty-seven *bad*/*good* Coordinated Antonymy database sentences, only two feature *bad* before *good*. However, a brief glance at these two examples (sentences 77b and 77c) shows that neither employ a standard *X and/or Y* framework. Rather, both sentences use the construction *X n as well as Y {n}*. This is relevant because it pushes the antonyms further apart, allowing the distance between X and Y to be closer to the distance created in Ancillary and Transitional Antonymy sentences. Perhaps this is why the irregular sequence of *bad* and *good* is not something which the writer of these sentences has sought to avoid. One might also argue that the semantics of *X as well as Y* places more emphasis on X than Y: sentence 77b refers to *the bad art as well as the good* and sentence 77c refers to *bad things as well as good*, making *bad* the 'surprise antonym' in both contexts. For these reasons, reverse sequence is less unacceptable in the framework *X as well as Y* than in the framework *X and/or Y*.

Therefore, a brief analysis of the seventeen reverse sequence *good*/*bad* sentences shows that syntactic distance is all-important. The closer together a pair of antonyms are in a given context, the greater the likelihood that they will conform to their normal sequence. When members of an antonymous pair are separated by more than one or two words in text, they become more likely

to reverse their normal sequence. However, even when uncoordinated, *good/bad* sentences favour normal sequence over reverse sequence at a ratio of nearly 5:1.

Morphological antonyms in reverse

One of the most important criteria when determining antonym sequence was deemed to be morphology; the nine morphologically related antonymous pairs in the database adhered to their favoured sequence most religiously. The class of antonymy which evidences word sequence bias most commonly is Coordinated Antonymy, especially when antonymous pairs are linked minimally by either *and* or *{n}or*. Therefore, this section will now examine the most 'anarchic' of examples – morphological word pairs linked only by *and* or *{n}or* which reverse their normal sequence.

The database of 3,000 sentences is too small to examine this phenomenon fully so the statistics below are based on the entire, 12-million-sentence corpus. Table 8.3 shows the nine morphological word pairs examined, the number of corpus sentences in which they occur in their normal sequence, and the number of corpus sentences in which they occur in their reverse sequence. This survey is limited to the frameworks *X and Y*, *X or Y* and *X nor Y* and therefore identifies word pairs which function in terms of Coordinated Antonymy and, to a lesser extent, Distinguished Antonymy (e.g. *the difference between X and Y*).

The figures presented in Table 8.3 are very convincing. In total, the nine antonymous pairs appear in an *X and/{n}or Y* environment in 590 sentences. In 584 of those sentences, the root word appears before the morphological antonym. Thus, numerous examples arise such as:

78a None, bar Germany, has experienced a flood of newcomers, though all have experienced a rise in immigration – **legal** and **illegal** – and have tightened their controls as a result, even without having passport controls on their national borders.

78b He said that whether the figure was **correct** or **incorrect** was now immaterial.

78c I can neither **prove** nor **disprove** the idea that the entire world is of my own creation: that not only this review, but also the book I'm reviewing, are products of my imagination.

In sentences 78a–c, the root word is the first-mentioned antonym and its morphologically derived partner is second-mentioned. In 99 per cent of *X and/or Y* contexts, this sequence is observed, even though reversal would entail no loss of grammaticality or shift in semantics. Therefore, although the sequence rule of antonymy is unspoken, it is also virtually unbroken by morphological pairs.

However, 'virtually unbroken' is not the same as 'totally unbroken' and six corpus sentences do reverse normal sequence, all of which are presented below.

Table 8.3 Sequence statistics for morphological antonyms in *X and/{n}or Y* framework in the corpus

		Normal sequence	*Reverse sequence*	*Total*
advantage	disadvantage	12	0	12
agree	disagree	48	1	49
correct	incorrect	13	0	13
directly	indirectly	365	3	368
honest	dishonest	3	0	3
legal	illegal	80	2	82
married	unmarried	25	0	25
officially	unofficially	9	0	9
prove	disprove	29	0	29
Total		**584**	**6**	**590**

The first triplet of sentences contain *directly* and *indirectly*. The reason that this pair yield as many as three rogue sentences is mostly attributable to their high raw frequency: in total, 368 *directly*/*indirectly* sentences follow an *X and/{n}or Y* framework, more than the other eight word pairs put together. The next two sentences feature *legal* and *illegal* in reverse sequence; the final sentence features *agree* and *disagree* in reverse sequence.

79a In my seventeen years with the group, I served **indirectly** and **directly**, no fewer than twenty chairmen and managing directors.
79b Through these holdings, CGIP **indirectly** and **directly** controls 20 per cent of CGS.
79c Economists estimate that between 1.5 and 2.5 jobs depend, **indirectly** or **directly**, on every Californian defence job.

80a The stricter approach coincides with growing demands for curbs from Congress, where bills have been tabled aimed at reducing both **illegal** and **legal** immigrants.
80b There is a thin line between **illegal** and **legal** drugs (that is, alcohol and tobacco).
80c There were 248 statements to **disagree** or **agree** with more or less strongly.

So, what do the above sentences have in common which allows them to present their antonyms in the 'wrong' sequence? This is not easy to judge. By a ratio of 120:1, one usually expects to find *directly and/or indirectly*, but the first three sentences above all refer to *indirectly and/or directly*. One possible explanation for this reversal is that the *indirectly* of each sentence is more contextually significant than the *directly*. For example, in sentence 79a, it is possible that the twenty chairmen served were mostly served *indirectly*; similarly, in sentence 79b, it is possible that the 20 per cent of CGS owned by CGIP is mostly held *indirectly*; and in sentence 79c it is possible that the jobs are mostly *indirectly* dependent.

Sentence 80a, which refers to *illegal and legal immigrants*, may be speculated about with more confidence. The phrase *illegal immigrants* is much more common that the phrase *legal immigrants* in text, occurring 691 times in the corpus, compared to twenty-eight occurrences of the latter. Therefore, the writer of this sentence may have first-mentioned *illegal* because it is more intuitively available than *legal*. This could help explain the reverse sequence, and the fact that this is one of only two corpus sentences which refer to *illegal and legal* rather than *legal and illegal*.

The other example is sentence 80b, which refers to *a thin line between illegal and legal drugs*. One explanation for this reverse sequence may lie in the post-modification of *legal drugs*. In parentheses, we find *that is, alcohol and tobacco*. This relates to *legal drugs*, so, had *legal drugs* appeared first in the sentence, this post-modification would have split the antonymous pair. Although it has been argued that identical syllables like to be separated in text, an entire interpolation may have been considered excessive. This may account for the reverse sequence of *legal* and *illegal*, although it should also be noted that *drugs*, like *immigrants*, collocates more with *illegal* than *legal*. Indeed, sentences 80a and 80b suggest that, in the framework *X and/or Y n*, the adjectival antonym which collocates better with the noun head is more likely to occupy X-position (or, to provide a weaker version of this hypothesis, these sentences suggest that collocational factors may explain why sequence rules are occasionally disregarded).

Sentence 80c is more difficult to explain, but it is possible that the statements in question showed a greater tendency to be disagreed with than agreed with. Whatever the logic behind this sequence, and the other five contexts above, such examples are very much the exception to the norm. Among the morphological antonyms chosen for study, normal sequence is adhered to in 99 per cent of sentences.

Lexical antonyms in reverse

In general, lexical antonyms are less prone to observe a given sequence than morphological antonyms. However, some pairs, such as *rightly/wrongly* and *confirm/deny*, have developed an unusually fixed sequence. In total, *rightly* and *wrongly* appear in an *X and/{n}or Y* framework in 170 corpus sentences. Of those 170 corpus sentences, *rightly* precedes *wrongly* in every one without exception.[13] In total, *confirm* and *deny* appear in an *X and/{n}or Y* framework in 317 corpus sentences. Of those 317 corpus sentences, *confirm* precedes *deny* in 315. The two rule-breakers are below:

81a When asked to **deny** or **confirm** whether he would be going to South Africa, the Bath centre said: 'Why do you want to know?'

81b Mr Sununu, 49, would neither **deny** nor **confirm** the reports, saying that 'a request isn't a request until the tall, thin guy sings', referring to Mr Bush.

There is nothing extraordinary about these sentences, although one could argue that the emphasis is more on denial than confirmation in each example because both sentences are accusatory. Perhaps the writer has reversed normal sequence to draw attention to the fact that neither respondent is actively denying their respective charge. One flaw in this explanation is that many *confirm/deny* sentences could be described in this way, yet only two corpus sentences in 317 place *deny* in front of *confirm*. This suggests that the rule of sequence usually over-rides any rhetorical consideration in the case of the lexical pairs *rightly/wrongly* and *confirm/deny*.

Antonyms: irreversible or idiomatic?

The majority of antonymous pairs in the database prefer one sequence over the other. In the case of some antonymous pairs (*rightly/wrongly*, *prove/disprove*, *correct/incorrect*), no evidence of that sequence being intra-sententially reversed can be found in the corpus. Once a sequence is established, it tends to become fixed, but many different criteria can influence why a particular antonymous pair favours a particular sequence. This chapter has shown that, where applicable, morphological pairs tend to occur with their root word first and their morphological antonym second, and that pairs with one positive and one negative term tend to mention their positive term first. Other factors such as magnitude, chronology, gender, phonology and idiomaticity can also affect antonym sequence.

However, the rules of sequencing are not as rigid as the rules of grammar and it is interesting to examine those sentences which flout convention. Sometimes, there would appear to be a specific reason for reversing normal sequence (an ancillary function may be being served or some rhetorical effect may be sought), but often no clear explanation can be found. It would seem that rules are there to be broken as well as obeyed.

This indicates that, technically, antonymous pairs should not be regarded as irreversible binomials because they are not irreversible in the same way as, say, 'foot and mouth disease' could never be known as 'mouth and foot disease'. However, to apply so strict a definition seems pedantic and, when coordinated, some antonymous pairs favour one sequence so strongly that it would be churlish not to regard them as irreversible binomials.

A more difficult issue involves whether or not these pairs should be regarded as idiomatic. In my view, the fact that most antonyms occasionally reverse their normal sequence, even when coordinated, prevents them from being truly idiomatic. These pairs are not 'unanalysable wholes' (McCarthy 1998: 148) like 'home and hosed' or 'huff and puff' and do not belong in the class of Idiomatic Antonymy alongside other fixed expressions.

Finally, to return to the altar, has this analysis shed any light on why wedding vows are taken 'for better for worse, for richer for poorer, in sickness and in health'? Well, morphology is not a factor here because all three pairs are lexical antonyms. But positivity does account for the appearance of *better*

before *worse*, and *richer* before *poorer*. Unfortunately, the final pair cast doubt on this explanation – why is *sickness* preferable to *health?* Perhaps phonological factors are also at work. Database evidence shows that shorter antonyms tend to precede longer antonyms, but this theory is also scuppered by the final pair because *sickness* has more syllables than *health*. However, as *better* also has more syllables than *worse*, it is possible that a rhythm is being maintained here. Maybe *better* precedes *worse* because it the more positive antonym, then *richer* precedes *poorer* for the same reason, and then *sickness* precedes *health* to provide some sort of phonological consistency. This explanation is not entirely satisfactory, but regardless of which hypothesis one prefers, it is clear that any rules of antonym sequence are susceptible to awkward counter-examples.

9 Antonymy, word class and gradability

When selecting a sample of antonymous pairs suitable for analysis, two of the criteria followed were that:

- word pairs selected should not be restricted to a single word class, but should feature nouns, verbs and adverbs, in addition to adjectives;
- words selected should not be restricted to one traditional category, but should include both gradable and non-gradable antonyms.

The reason for the first criterion is that the phenomenon of antonymy is not restricted to a single word class, even though early analyses of the sense relation tended to focus primarily on adjectives (e.g. Lyons 1977). However, just as a majority of native speakers would identify *heavy* as being the 'opposite' of *light*, so too would a majority of native speakers identify *strength* as being the 'opposite' of *weakness*, *win* as being the 'opposite' of *lose*, and *quickly* as being the 'opposite' of *slowly*. Therefore, antonymous nouns, verbs and adverbs were included in the database as well as adjectives.

The reason for the second criterion is that, if asked to provide a list of 'opposites', most native speakers would not discriminate between gradable and non-gradable pairs; antonyms such as *new*/*old*, *happy*/*sad* and *poor*/*rich* would be interspersed with antonyms such as *female*/*male*, *alive*/*dead* and *false*/*true*. Despite this, Cruse (1986) and Lyons (1977) do not regard non-gradable pairs as 'antonyms' at all, reserving that label for gradable pairs. This seems counter-intuitive and the database was accordingly designed to incorporate both gradable and non-gradable antonymous pairs.

This chapter will now consider two questions: whether the function of antonymy in text is related to word class and whether the function of antonymy in text is related to gradability.

Does word class affect the function of antonymy?

Fellbaum suggests that there is 'nothing special' (1995: 285) about adjectival antonyms and this section will report on the extent to which evidence from a large corpus substantiates this claim. In order for the relationship between

word class and the function of antonymy to be assessed, the original database of 3,000 sentences needed to be broken down according to whether the antonymous pair were adjectives, nouns, verbs or adverbs. This caused some headaches because a number of antonyms are able to cross word class. For example, 57 per cent of *poor/rich* database sentences feature those words operating as nouns,[1] but the remainder feature those words operating as adjectives. On the other hand, some antonymous pairs which can act as both adjectives and nouns favour the former word class: *right/wrong* functions adjectivally at a rate of 82 per cent; *new/old* at 86 per cent; *private/public* at 81 per cent; and *old/young* at 75 per cent. An unusual pair is *hate/love* because they occur both as nouns and verbs, with the distribution favouring the latter at a rate of 82 per cent.

The *un*-words component of the sample breaks down as follows: 112 of the 156 sentences sampled feature adjectival antonyms; 16 feature nouns; 20 feature verbs and 8 feature adverbs. This is interesting because these pairs have not been chosen according to any grammatical criteria; the only prerequisite is that they are morphologically linked by the prefix *un*. Therefore, from this admittedly small sample, we could infer that about 71 per cent of *un*-antonyms are adjectives, about 11 per cent are verbs, about 13 per cent nouns and about 5 per cent adverbs. Whether this distribution would be similar for non-morphologically related antonyms is nearly impossible to gauge, though it seems fair to say that most antonyms in English are probably adjectival. It should also be noted that this distribution would vary according to the morpheme used to retrieve sentences: for example, antonyms formed by the prefix *dis-* would show a stronger bias towards verbal pairs (*like/dislike*, *connect/disconnect*, *appear/disappear*, etc.).

In total, the 3,000 database sentences feature 1,739 instances of adjectival antonymy, 520 instances of nominal antonymy, 427 instances of verbal antonymy, and 314 instances of adverbial antonymy. Table 9.1 presents the distribution of these sentences in relation to the new classes of antonymy to which they have been assigned. This table shows that, of the 1,162 sentences belonging to the category of Ancillary Antonymy, 704 are adjectives, 161 are nouns, 185 are verbs and 112 are adverbs. Or, to look at the data from another perspective, of the 427 verbal pairs in the database, 185 function in an ancillary fashion, 206 function in a coordinated fashion, 25 are comparative, 8 are negated and 3 are idiomatic.

These figures are expressed as a normal distribution score in Table 9.2.

Table 9.1 Distribution of antonymy by word class (raw frequency)

	An	*Co*	*Cm*	*Ds*	*Tr*	*Ng*	*Ex*	*Id*	*Oth*	*Total*
Adjectives	704	636	112	93	58	23	27	20	66	**1,739**
Nouns	161	129	55	68	32	26	9	0	40	**520**
Verbs	185	206	25	0	0	8	0	3	0	**427**
Adverbs	112	180	13	0	0	5	4	0	0	**314**
Total	**1,162**	**1,151**	**205**	**161**	**90**	**62**	**40**	**23**	**106**	**3,000**

Table 9.2 Distribution of antonymy by word class (normal distribution score)

	An	*Co*	*Cm*	*Ds*	*Tr*	*Ng*	*Ex*	*Id*	*Oth*
Adjectives	0.70	0.36	0.45	0.55	0.63	0.26	0.70	0.89	0.60
Nouns	0.11	0.12	0.92	0.91	0.90	0.93	0.78	0.21	0.91
Verbs	0.85	0.68	0.37	0.23	0.21	0.37	0.07	0.66	0.22
Adverbs	0.36	0.87	0.17	0.23	0.21	0.31	0.57	0.21	0.22

A score above 0.5 indicates that a given function of antonymy is favoured by antonyms belonging to a given word class; a score below 0.5 indicates that a given function of antonymy is disfavoured by antonyms belonging to a given word class. For example, adverbial antonyms are disproportionately inclined towards the class of Coordinated Antonymy (normal distribution score: 0.87), but are less likely to operate in terms of Ancillary Antonymy (0.36). In relative terms, adverbial antonyms can also be seen to shy away from all minor classes of antonymy except Extreme Antonymy, to which they show a slight proclivity (0.57). The most unusual word class distribution belongs to nouns because they serve minor functions of antonymy at proportionally higher rates than pairs which belong to other word classes. Five of the six minor classes are over-represented by nouns, suggesting that nominal antonyms are less inclined towards the two dominant classes and, therefore, are less textually predictable. However, none of the normal distribution scores recorded in Table 9.2 are statistically significant (i.e. higher than 0.95 or lower than 0.05), which indicates that the connection between word class and antonym function is not as strong as one might expect. All eight new classes of antonymy will now be briefly re-examined to determine the extent to which they are able to accommodate adjectival, nominal, verbal and adverbial pairs.

Ancillary Antonymy

In total, 38.8 per cent of database sentences feature Ancillary Antonymy. And if we examine these sentences word class by word class, we find relatively little variation from this figure: among adjectives, the proportion rises slightly to 40.5 per cent; among nouns, it drops to 31.0 per cent; and among verbs and adverbs, it is 43.3 per cent and 35.7 per cent respectively. As none of these proportions differ dramatically from the average for all antonyms, Ancillary Antonymy would seem able to cross word class quite easily, paying little attention to whether the antonymous pair are adjectives, nouns, verbs or adverbs. The following four sentences demonstrate this:

82a Cane is a victim of what Ben Fletcher, professor of business psychology and Dean of the Business School at the University of Hertfordshire, defines as the classic stressful situation – a job with **high** *demands* but **low** *support*.

82b However, it is the scale of *Labour* **success**, not of *Conservative* **failure** that stands out.

82c The key question was this: at what point does *sport* **end** and *political manipulation* **begin**?

82d The mood at Labour's headquarters, a redundant school in Chepstow, was **publicly** *restrained* and **privately** *buoyant*.

Although the quartet of sentences above are not alike in every way, they each exemplify the phenomenon of Ancillary Antonymy. As discussed in Chapter 4, an antonymous pair is drawing out an opposition between a pair of words or phrases which may not otherwise be considered contrastively. For example, in sentence 82a, the antonymous pair of adjectives, *high* and *low,* highlight the contrast between the nouns *demands* and *support*; in sentence 82b, the established antonyms are the noun heads *success* and *failure* which signal an opposition between *Labour* and *Conservative*; sentence 82c relies on an antonymous verbal pair (*begin*/*end*) to create a distinction between the noun phrases *sport* and *political manipulation*; and sentence 82d uses a pair of antonymous adverbs (*privately*/*publicly*) to enhance the contrastive power of *restrained* and *buoyant*.

In other words, antonyms are being used to signal contrast regardless of which of the four word classes they belong to. Furthermore, B-pairs also seem to cross the boundaries of word class, being nouns in sentences 82a–c and being adjectives in sentence 82d. Table 9.2 demonstrates that Ancillary Antonymy is more popular among verbs and adjectives than it is among adverbs and nouns, but as pairs belonging to all four word classes function this way at a rate of at least 30 per cent, it would seem that the antonymity of a given pair is more important in an Ancillary Antonymy sentence than its word class.

Coordinated Antonymy

In total, 38.4 per cent of database sentences were assigned to the class of Coordinated Antonymy. Among adjectives, the proportion is very similar (36.6 per cent), but among other word classes, differences arise: nouns disfavour Coordinated Antonymy (24.8 per cent), but verbs and adverbs favour the class (48.2 per cent and 57.3 per cent).[2] One could interpret this distribution as evidence that the word class of an antonymous pair sometimes influences its textual function. However, one could also argue that the word class distribution of coordinated antonyms still reflects similarity, albeit a lesser similarity than those shown by ancillary antonyms. For example, one can safely state that the phenomenon of Coordinated Antonymy pervades all four word classes. Among nouns it is least common, yet one-quarter of database sentences (129 out of 520) which feature antonymous nouns still place that pair in a coordinated framework. To illustrate the ability of coordinated antonyms to cross word class, two frameworks closely associated with this class (*X as well as Y* and *both X and Y*) will be exemplified.

X as well as Y

83a It would be interesting to hear all experiences, **good** as well as **bad**.
83b John Hoddinott, president of Acpo and Chief Constable of Hampshire, said DNA testing is a tremendously powerful tool – it proves **innocence** as well as **guilt**.
83c Part of the fuss can be explained by our ambivalent feelings about the Duchess of York, whom the press, and perhaps even their readers, soon learned to **hate** as well as **love**.
83d John Major has repeatedly said, **privately** as well as **publicly**, that I have his complete backing.

The four database sentences above show that antonyms can function as part of the same lexico-syntactic framework regardless of word class. Here, the pattern *X as well as Y* occurs with antonymous adjectives (*good*/*bad*), nouns (*innocence*/*guilt*), verbs (*hate*/*love*) and adverbs (*privately*/*publicly*). Though the immediate environment of each antonymous pair remains unaffected by word class, the four examples are not identical. For instance, the phrases *good as well as bad* and *privately as well as publicly* are entirely removable from their context, whereas *innocence as well as guilt* and *hate as well as love* are not removable. However, nouns and verbs are more essential in syntax than adjectives and adverbs, which are often removable in text whether antonymous or not. The important fact is that the function of *X as well as Y* is similar in each example: regardless of word class, antonyms always express an encompassing quality when inserted into this framework.

both X and Y

84a In this city, at least, it is as funky to be passionate about films – both **old** and **new** – as it is to be about football everywhere else.
84b Like all of us, athletes need to find a way of rationalising both **failure** and **success**.
84c When I hear grateful tourists saying their meal was 'very good', I both **agree** and **disagree**.
84d Unless we can secure huge changes in attitudes to drug abuse, many thousands will continue to suffer both **directly** and **indirectly** from the consequences.

The antonymous pairs of sentences 84a–d each occur in the same framework (*both X and Y*) and serve a similar role (to signal inclusiveness), despite belonging to different word classes. This confirms that frameworks typical of Coordinated Antonymy are able to accommodate antonyms expressed by adjectives, nouns, verbs and adverbs. However, that is not to say that the likelihood of an antonymous pair of nouns occurring in a coordinated environment is the same as the likelihood of a pair of antonymous adverbs

occurring in the same environment. The fact that adverbs and verbs tend to favour Coordinated Antonymy at double the rate of nouns is indicative of the unusual distribution of nominal antonyms across new classes. Table 9.1 demonstrates that Ancillary Antonymy and Coordinated Antonymy remain the most popular classes among antonymous nouns, but Table 9.2 indicates that, in relative terms, these two dominant classes of antonymy are shunned by nouns in favour of lower-frequency functions.

Comparative Antonymy

As Table 9.1 shows, Comparative Antonymy is another class of antonymy which is able to cross the boundaries of all four word classes. In total, 6.8 per cent of database sentences were attributed to the class of Comparative Antonymy. Among adjectives, verbs and adverbs, proportions remain fairly constant (6.4 per cent, 5.9 per cent and 4.1 per cent respectively), but among nouns, this figure rises to 10.6 per cent, generating a normal distribution score of 0.92. In practice, this means that the database contains many examples of, say, *punishment* being compared with *reward* (21.1 per cent of sentences featuring this pair of words were assigned to Comparative Antonymy: treble the average for all pairs) and *strength* being compared with *weakness* (17.1 per cent), but fewer examples of adjectives, verbs and (especially) adverbs functioning in this way. Four of the 205 sentences retrieved and classified in terms of Comparative Antonymy are listed below, one representing each of the word classes analysed:

85a She objected to a system whereby if you are **rich** you can buy access to programmes more tasteless and more indecent than if you are **poor**.

85b Institutional investors may, however, be sympathetic in special cases as they realise that the retention of key executives is as, if not more, important in a **recession** as in a **boom**.

85c Douglas Hurd, the Foreign Secretary, said afterwards: 'It's always better to **win** than to **lose**, but it doesn't affect our ability to ratify the treaty.'

85d In every part of the country, more people think **badly** of him than think **well**.

The quartet of sentences above do not adhere tightly to a strict lexico-syntactic framework like those associated with Coordinated Antonymy, but they each feature a comparison between a pair of antonymous concepts: sentence 85a compares *rich* and *poor* in terms of access to programmes; sentence 85b compares *recession* and *boom* in terms of the importance of retaining executives; sentence 85c compares *win* and *lose* in terms of quality; and sentence 85d compares *badly* and *well* in terms of how people think of someone. Word class seems to hold relatively little sway over the semantic function served by

antonymy in these examples. Indeed, one could easily imagine a nominalised version of *rich* and *poor* in sentence 85a ('the rich can buy access to programmes more tasteless and more indecent than the poor'), and *win* and *lose* in sentence 85c ('winning is always better than losing'). This word class flexibility confirms that the antonymous pairs in the sentences above appear to have been chosen more because of their semantic opposition than because of any grammatical criteria.

However, once again, this is not to imply that word class is an irrelevance; the fact remains that, according to this research, a pair of antonymous nouns has double the chance of functioning in a comparative context than a pair of antonymous adverbs. This illustrates that although most functions of antonymy can be served by word pairs belonging to all four word classes, variation between individual classes is not uncommon.

Negated Antonymy

So far, the three classes of antonymy examined have each been available to adjectives, nouns, verbs and adverbs. This is also true of Negated Antonymy, even though the total number of database sentences attributed to this class is much lower (62, compared with 1,162 ancillary contexts, 1,151 coordinated contexts and 205 comparative contexts). In total, 2.1 per cent of database sentences were assigned to Negated Antonymy: antonymous nouns (5.0 per cent) favour the class most strongly, but adjectives (1.3 per cent), verbs (1.9 per cent) and adverbs (1.6 per cent) also occur in negated frameworks, as the following sentences illustrate:

86a Here Farrell's complicity becomes **active**, not **passive**, as he promotes an architecture (Post-Modernism) suggesting that you can have it both ways: that an architecture of facades need not be compromised by whatever it contains.

86b Sponsors want to invest in **success**, not **failure**.

86c You purport to be a national newspaper, not an extremist group like the Hackney Community Defence Association (HCDA): H – they are not Hackney people, C – they are not community-based people, D – they **attack**, not **defend**, A – they are a narrow, highly secretive group, not open and fair.

86d Somewhere between Everything You Wanted To Know About Sex (But Were Afraid To Ask) and Annie Hall, he mastered the ability to have actors act **well**, not **badly**.

The four examples above provide further evidence that antonymous pairs serve certain functions in text and that these functions are not bound by the restrictions of word class. Each sentence places antonyms within the framework *X not Y*, negating the second antonym in order to bolster the first. This

strategy can be seen in operation with adjectives in sentence 86a, nouns in sentence 86b, verbs in sentence 86c and adverbs in sentence 86d. Though, like most minor classes, Negated Antonymy seems especially popular with antonymous nouns, it is able to accommodate all four word classes under scrutiny and yields a fairly even distribution despite its relatively low frequency.

Distinguished Antonymy

Even though the class of Distinguished Antonymy accounts for over twice as many database sentences as Negated Antonymy, it does not feature sentences which belong to all four word classes. In total, just over 5 per cent of all database sentences were assigned to this class. However, all 161 of these sentences feature antonyms which belong to one of two word classes: adjectives and nouns. The reason for this is that one can distinguish between 'things' more easily than one can distinguish between 'actions'. Thus, pairs of nouns and noun modifiers (adjectives) are differentiated between more frequently than pairs of verbs and verb modifiers (adverbs). Further proof of this is provided by the word class distribution of Distinguished Antonymy sentences: Table 9.2 shows that adjectives record a normal distribution score of 0.55, but nouns record a more extreme score of 0.91, confirming the proportional bias towards 'things'. The first triplet of sentences below show Distinguished Antonymy sentences which feature antonymous adjectives; the second show Distinguished Antonymy sentences which feature antonymous nouns.

87a This punishment did not discriminate between the **active** and **passive** conspirators but nor did it alter the fact that all alike had been let off lightly in the courts.
87b But one of the failings of the play is that it never allows for any distinction between **true** and **false** rhetoric.
87c More precisely than he knew, Mr Todd, of the Transport and General Workers Union, defined a crucial difference between the **old** and **new** approaches to Labour politics.

88a To them, unlike Mr Bragg, the distinctions between **fact** and **fiction** or **good** and **bad** are not always so obvious.
88b It's true that his advocacy can make the difference between **success** and **failure** for a young band.
88c But then, crowds do not discriminate any too nicely between **guilt** and **innocence** when their blood is up.

All six of the sentences above differentiate between a pair of antonymous noun phrases. In sentences 87a–c, the antonymy is expressed adjectivally, enabling contrast to arise between different aspects of the same concept (namely *active and passive conspirators*, *true and false rhetoric* and *old and new approaches to Labour*

politics). In sentences 88a–c, the antonymy is expressed by nouns, enabling contrast to arise directly between different concepts (namely *fact and fiction*, *good and bad*, *success and failure* and *guilt and innocence*).

The typical lexico-syntactic framework associated with Distinguished Antonymy (*difference between X and Y*) is not popular among verbs (unless nominalised, i.e. *winning/losing*) and nor is it popular among adverbs. The sentence below shows that it is possible for adverbs to be differentiated in text, but this example was retrieved not from the database, but from elsewhere in the corpus.

89a The difference between a **well** and **badly** structured fund can be startling.

This sentence is syntactically akin to the previous six in that two antonymous noun phrases, linked by *and*, follow the words *difference* and *between*. The only difference is that antonymy here is expressed not by adjectives or nouns, but by adverbs. Database evidence suggests that such constructions are very rare, though the example above does prove that Distinguished Antonymy is not always restricted to adjectives and nouns.

Transitional Antonymy

The class of Transitional Antonymy is similar to Distinguished Antonymy because it too is occupied exclusively by adjectives and nouns, in the database at least. Table 9.2 shows that antonymous adjectives achieve a normal distribution score of 0.63 and antonymous nouns of 0.90, making the word class distribution of Transitional Antonymy almost identical to that of Distinguished Antonymy: nouns are favoured, but verbs and adverbs are strongly disfavoured. And just as the statistical distribution of these two classes is similar, so the explanation is similar: Transitional Antonymy sentences describe a movement from one antonymous state to another, and this state tends to be a noun phrase with antonymy expressed adjectivally or via the noun head:

90a How easy to slip from the **legal** to the **illegal** trade, especially when the law is so patchy and the temptation so great.
90b The main problem with the proposal, as with all radical reforms, is the difficulty of transition from the **old** system to the **new**.
90c How does he effect the transition from **private** to **public** self, like Wonderwoman going into a spin?

91a This would undermine the UN's hope that a tidy transition from **war** to **peace** would set a precedent for UN attempts to defuse crises elsewhere.
91b The report does highlight the possibility of a new golden age, where developing countries have the opportunity to export services to the

developed world, though it does not really explore some of the more interesting issues raised by falling communication costs, such as whether this means that countries which have suffered from peripheral geographic location can turn **disadvantage** to **advantage**.

91c But it's been strangely quiet in Twickenham over the last week or two, and I must say my **optimism** is turning to **pessimism**.

The transition in each of the above sentences is from one noun phrase to another. In the case of the first triplet of sentences, antonymy arises in the pre-modification of the noun head (e.g. *from the legal to the illegal trade*); in the case of the second triplet of sentences, antonymy arises within the noun head (e.g. *from war to peace*). As with Distinguished Antonymy sentences, verbs and adverbs are less easily accommodated by Transitional Antonymy frameworks and no examples of either were found in the database. However, verbs and adverbs are not precluded by definition from this category – it is possible to construct artificial contexts in which members of these word classes form the basis of a contrast ('they have gone from winning elections to losing elections'; 'she went from playing well to playing badly').

Extreme Antonymy

Even among low-frequency classes of antonymy, equal distribution across word class is not uncommon. For example, despite only forty database sentences being assigned to the class, Extreme Antonymy shows a very similar dispersion between adjectives, nouns and adverbs (1.6 per cent, 1.7 per cent and 1.3 per cent respectively), although the database records no example of antonymous verbs functioning in this way. Examples of the three word classes assigned to Extreme Antonymy are recorded below:

92a Derek Bluston, of the solicitors Malkin Janners, says: 'To get involved in litigation you either have to be very **rich** or very **poor**.'

92b Reflecting on recent events, he said: 'It wasn't a total **success** but it wasn't a total **failure**.'

92c Mr Baker rejected 'armchair critics' who had attacked the police for either responding too **quickly** or too **slowly** to street disturbances.

Idiomatic Antonymy

In the database, idiomaticity happens only to occur with antonymous adjectives and verbs (*to blow hot and cold*, *to agree to disagree*, etc.). However, a wider sample of antonymous pairs may well have also retrieved idiomatic contexts which feature antonymous nouns and adverbs. Therefore, the fact that only two word classes have been classified in terms of Idiomatic Antonymy may not be of significance.

Overview of word class and antonymy

This analysis of word class suggests that the use of antonymy is influenced more by semantic than grammatical factors. That is not to say no correlation at all emerges between textual functions of antonymy and word class, but this correlation is not as strong or as consistent as one might expect. In the case of all four word classes examined, at least 55 per cent of database sentences fall into one of the two major classes, Ancillary and Coordinated Antonymy. Adverbs tend to favour the latter class and verbs tend to favour the former class, but these preferences are slight. What is important is that these classes are not simply the product of grammatical factors; it is not true, for instance, that all Coordinated Antonymy sentences feature adjectives or that all Ancillary Antonymy sentences feature verbs.

On the other hand, some new classes of antonymy avoid certain word classes entirely because it is grammatically difficult (or even impossible) to house these words within their associated frameworks. Thus, verbal and adverbial antonyms are not found in any Distinguished or Transitional Antonymy database sentences. However, not all minor classes discriminate in this way and the word class distribution of Negated Antonymy, for example, is remarkably consistent.

Indeed, although the statistical analysis presented in Table 9.1 is open to interpretation in many ways, I do not believe that word class affects the textual function of antonymy to any great extent. Writers use antonyms to much the same ends, regardless of whether those antonyms are adjectives, nouns, adverbs or verbs. Therefore, antonymy is not only a phenomenon which crosses word classes, it is (to some degree, and especially in the case of the two major classes) a phenomenon which functions the same irrespective of word class. This is no great surprise when one considers how similar sentiments can be expressed by antonyms belonging to different word classes:

- Will her application be **successful** or **unsuccessful**?
- Will her application be a **success** or a **failure**?
- Will her application **succeed** or **fail**?
- Will she apply **successfully** or **unsuccessfully**?

The four questions above[3] are not semantically identical, but they are very similar. Antonymy is expressed by adjectives, nouns, verbs and adverbs respectively, but only minor grammatical adjustments are necessary to accommodate these different word classes. Although the verb *to succeed* can morph into *success*, *successful* and *successfully*, the verb *to fail* is less flexible. The noun form, *failure*, is no problem, but the corresponding adjective and adverb require a change of root word in order for *unsuccessful* and *unsuccessfully* to be derived. However, all four sentences still feature a recognisable antonymous pair and the grammatical surgery required to change the context according to the word class of each antonymous pair is largely cosmetic. In other words, these four questions

illustrate that antonyms can serve the same textual function (in this case, Coordinated Antonymy) regardless of whether they are adjectives, nouns, verbs or adverbs. Fellbaum talks about 'the arbitrariness of the encoding of concepts into particular word classes' (1995: 285) and this study confirms that antonyms may be encoded in a variety of grammatical forms, but will continue to serve the same textual functions at similar rates. Antonymity is more important than word class.

Does gradability affect the function of antonymy?

Table 9.1 shows that the database contains 1,739 sentences which feature adjectival antonyms. These sentences have been further broken down according to whether the antonymous pair therein is gradable or non-gradable. The various classifications of antonymy developed by semantic theorists such as Leech (1974), Lyons (1977) and Cruse (1986) indicate that this is the most fundamental logical distinction which can be made between adjectival antonymous pairs. Word pairs which refer to either end of a scale but leave 'semantic space' between themselves are known as gradable antonyms; word pairs which collectively exhaust all 'semantic space' in their scale are known as non-gradable antonyms. Examples of the former include *cold* and *hot*, which sandwich terms such as *warm*, *tepid*, *chilly* and *cool*; examples of the latter include *alive* and *dead*, which leave no room for in-between descriptions – one is either living or not living. But does the status of an antonymous pair affect the way in which they are used in text?

Gradable and non-gradable pairs

The sample of fifty-six antonymous pairs includes thirty-one adjectival pairs,[4] of which twenty-five are gradable. The six non-gradable pairs are *alive/dead*, *correct/incorrect*, *false/true*, *female/male*, *illegal/legal* and *married/unmarried*. Paradoxically, the term 'non-gradable' does not mean that these words are never graded in text; one can easily imagine a context in which each of these words might be modified, most notably *true*, which is often preceded by *very*, *quite*, *extremely*, etc. The definition of non-gradable antonymy is that the application of one antonym logically precludes the application of the other. Put more simply, if something is currently *illegal* in the UK, it cannot simultaneously be *legal* in the UK. However, the distinction between gradable and non-gradable pairs is intuitive and often fine,[5] as Palmer points out when noting that 'some pairs of adjectives, e.g. *honest/dishonest*, *obedient/disobedient*, *open/shut* are . . . gradable in terms of more and less, yet . . . the denial of one is usually taken to assert the other' (1976: 81). This section will examine whether the gradable/non-gradable distinction is supported by textual evidence. If the function of adjectival antonymous pairs is related to whether they are gradable or non-gradable, then this logical distinction can be seen as an important factor which actively influences the way in which antonyms are stored and

implemented; however, if gradable pairs are used in much the same way as non-gradable pairs, the relevance of this traditional dichotomy will be further undermined.

Collectively, the six non-gradable pairs feature in 283 database sentences and the twenty-five gradable pairs feature in 1,456 database sentences. Their distribution across new classes of antonymy is recorded in Table 9.3. Immediately, one is struck by the fact that all eight new classes of antonymy, even the minor classes, feature adjectival antonyms which are non-gradable as well as adjectival antonyms which are gradable. It certainly could not be said that the logical distinction made about antonymous pairs by theoretical semanticists is reflected by textual evidence – the function of antonymy is not determined by the gradability of the pair.

However, some patterns do emerge in Table 9.3, such as the dissimilarity of frequency between non-gradable ancillary antonyms (69) and non-gradable coordinated antonyms (144). This is noticeable because the database distribution of antonymous pairs across these two classes is usually very similar (38.7 per cent and 38.4 per cent in total, though slightly more biased towards Ancillary Antonymy among adjectival pairs). This anomaly becomes more apparent if the raw frequency statistics of Table 9.3 are expressed in percentage terms.

Table 9.4 shows the proportion of gradable and non-gradable pairs which fall into each new class of antonymy. For example, 43.6 per cent of gradable pairs in the database were attributed to the class of Ancillary Antonymy, 33.8 per cent to the class of Coordinated Antonymy and 6.2 per cent to the class of Comparative Antonymy. These figures are very similar to the average for all adjectival antonyms of 40.5 per cent, 36.6 per cent and 6.4 per cent respectively, though this similarity is not unexpected given the dominance of gradable pairs overall. However, if the percentages recorded by non-gradable pairs are examined, more relevant deviations surface. For example, less than one-quarter of non-gradable pairs occur in Ancillary Antonymy sentences, but

Table 9.3 Distribution of gradable and non-gradable pairs across new classes (raw frequency)

	An	*Co*	*Cm*	*Ds*	*Tr*	*Ng*	*Ex*	*Id*	*Oth*	*Total*
Gradables	635	492	90	71	54	21	25	18	50	1,456
Non-gradables	69	144	22	22	4	2	2	2	16	283
All adjectives	**704**	**636**	**112**	**93**	**58**	**23**	**27**	**20**	**66**	**1,739**

Table 9.4 Distribution of gradable and non-gradable pairs across new classes (%)

	An	*Co*	*Cm*	*Ds*	*Tr*	*Ng*	*Ex*	*Id*	*Oth*	*Total*
Gradables	43.6	33.8	6.2	4.9	3.7	1.4	1.7	1.3	3.4	100
Non-gradables	24.4	50.9	7.8	7.8	1.4	0.7	0.7	0.7	5.6	100
All adjectives	40.5	36.6	6.4	5.3	3.3	1.4	1.6	1.2	3.7	100

over half occur in Coordinated Antonymy sentences. In other words, non-gradable antonyms show a marked preference for coordinated constructions and tend to favour sentences such as those below:

93a I doubt there are many men, **alive** or **dead**, who have not come a cropper, at one time or another, at the hands of a member of the UNfairer sex.
93b Bishop Lindsay carefully avoided any reference to the reason why Anglicans, **married** and **unmarried**, are joining the exodus to Rome.
93c Part of the novelty was that the masters of the flat and the crown codes, **male** and **female**, were brought together.

The trio of sentences above show non-gradable antonymous pairs serving to exhaust the scale against which they operate. Thus, the men of sentence 93a are said to be *alive or dead*; the Anglicans of sentence 93b are *married and unmarried*; and the masters of the flat and crown codes of sentence 93c are *male and female*. One reason why non-gradable antonyms tend to be used in this fashion proportionally more often than gradable antonyms is because non-gradable antonyms are better equipped to exhaust their entire scale. For evidence of this, compare the triplet of sentences above with the triplet of sentences below, each of which feature a pair of gradable antonyms.

94a The corner barber shop, a few streets away from my house, is a popular meeting place for Jamaican men, **young** and **old**.
94b Labour proposes new tax bands for **rich** and **poor**.
94c Because making Bills, **good** or **bad**, gets ministers on television and makes government look important.

Lexico-syntactically, sentences 94a–c are very similar to sentences 93a–c. They also serve a similar function – to signal inclusiveness or exhaustiveness of scale. However, it could be argued that gradable antonyms do not serve this function as effectively as non-gradable antonyms. For example, the phrase *men, old and young* may be synonymous with 'men of all ages', but it could equally refer exclusively to 'old men and young men', which does not exhaust an entire range. Similarly, the tax bands alluded to in sentence 94b could apply to *rich and poor* (meaning 'all people, regardless of wealth') or *rich and poor* (meaning 'those at either end of the wealth scale'). On the other hand, *dead or alive*, *married and unmarried* and *male and female* automatically exhaust their entire scale because no other points on that scale exist. Thus, non-gradable antonyms arise in proportionally more Coordinated Antonymy sentences because they are more semantically exhaustive. Gradable antonyms may signal exhaustiveness of scale, but non-gradable antonyms necessitate exhaustiveness of scale.

However, this raises a further question: why do writers choose to record one non-gradable pair at the expense of another? For example, to return to

sentence 93a, it would be equally feasible (or even more feasible) to describe the men as *married or unmarried*, rather than *alive or dead*. So why is the latter scale identified? One could even ask why any non-gradable scale should be identified given that one antonym or the other must always apply to the corresponding noun head. The answer to this question is that, in addition to signalling exhaustiveness of scale, the non-gradable pairs of sentences 93a–c fulfil another, perhaps more important role: they identify a relevant scale. For example, *married and unmarried* signals that this is the appropriate scale against which we should measure Anglicans in sentence 93b. The reason why an appropriate scale needs to be identified is not always apparent when glancing at out-of-context corpus sentences. However, the notion of the 'surprise antonym' may be relevant once again as sentences 93a and 93c are clearly weighted towards *dead* and *female* respectively.

Because non-gradable antonyms favour Coordinated Antonymy environments, other classes are inevitably disfavoured. Therefore, we find that the ratio of gradable to non-gradable antonyms in Ancillary Antonymy sentences is as high as 9:1, despite being nearer 3:1 among Coordinated Antonymy examples. This means that gradable pairs are used to signal contrast more frequently, in relative terms, than non-gradable pairs. Among low-frequency classes, proportions vary: the gradable/non-gradable ratio for Comparative Antonymy sentences is 4:1 and for Distinguished Antonymy sentences it is 3:1. This suggests that non-gradable pairs show a slight bias towards these classes, as the global database ratio for antonymous pairs is approximately 5:1. This is surprising as one would expect texts to differentiate between pairs such as *poor* and *rich* more frequently than pairs such as *dead* and *alive*, where the distinction is usually self-evident. Transitional opposition favours gradable pairs by a ratio of 13:1, Negated Antonymy by 10:1 and Extreme Antonymy by 12:1. Unfortunately, with fewer than a handful of sentences for each class featuring non-gradable antonyms, safe conclusions are not easily drawn.

However, it is reasonable to state that, in general, writers use non-gradable pairs in much the same way as they use gradable pairs. All eight new classes of antonymy identified yield examples of both gradable and non-gradable pairs and the majority of these classes display no strong bias either way. However, one generalisation may be made from the data: non-gradable antonyms do tend to favour the class of Coordinated Antonymy. Over half of all occurrences of *alive/dead*, *correct/incorrect*, *false/true*, *female/male*, *illegal/legal* and *married/unmarried* function in a coordinated framework, but only one-third of gradable adjectives are used in a similar way. From this, it would seem that the less gradable a pair of words are, the more likely they are to express exhaustiveness in text. However, this generalisation is based on an analysis of only six non-gradable pairs and their distribution across only 283 sentences. The most striking aspect of this data, in the view of this analyst at least, is the similarity of behaviour among gradable and non-gradable pairs, not the dissimilarity.[6]

Different pairs, similar usage

Antonyms are antonyms, regardless of word class, regardless of gradability. That is the (slightly overstated) conclusion of initial, corpus-based research into the ways in which different kinds of antonymous pairs function in text. First, the word class to which an antonymous pair belongs was examined to discover whether this influences the way that it is likely to operate in text. Data evidences some correlation, but this correlation is relatively insignificant; Coordinated Antonymy, Ancillary Antonymy, Comparative Antonymy and Negated Antonymy are each able to accommodate antonyms that belong to all four word classes investigated. Adverbs and verbs display a slight bias towards Coordinated Antonymy, while antonymous nouns tended to disfavour this class. One important proviso to this generalisation is that some new classes of antonymy did not accommodate certain word classes at all because of grammatical restrictions: no examples of adverbs or verbs being classified as Distinguished Antonymy or Transitional Antonymy were found in the database. This suggests that the textual profile of antonymous verbs and adverbs is slightly different from the textual profile of antonymous adjectives and nouns. However, in general, the function of antonymy remains oblivious to word class; when antonyms are used, semantics takes precedence over grammar.

This is confirmed by the second part of the chapter, which presented an analysis of whether the traditional category to which an antonymous pair had been assigned influences the functions that it is likely to serve in text. Thus, the adjectival component of the database was broken up according to whether the antonyms were gradable or non-gradable. It was found that non-gradable pairs showed a bias towards Coordinated Antonymy and served an ancillary role less often than their gradable counterparts. However, once again, it was striking to note how minor the differences are between the textual function of gradable and non-gradable antonyms. Non-gradable pairs occurred in sentences belonging to all eight new classes identified, and, in the case of the majority of classes, the distribution was relatively similar to that of gradable pairs.

Therefore, corpus evidence suggests that the textual function of antonymy is not greatly influenced by either word class or by gradability. Coordinated Antonymy is disproportionately favoured by verbs, adverbs and non-gradable adjectives, while antonymous nouns incline more towards the classes of Comparative, Distinguished, Transitional and Negated Antonymy. However, these proclivities, though not irrelevant, remain relatively insignificant given that traditional studies of antonymy often focused solely on adjectives and disputed whether non-gradable pairs were valid antonyms at all.

10 Tomorrow's antonyms

At the start of the twentieth century, *gay* and *straight* held no obvious semantic relation; in the middle of the twentieth century, the words remained unrelated; now, at the start of the twenty-first century, they would be readily identified as 'opposites'. But why do certain pairs of words have antonymous status conferred upon them? What does this process involve? And where do new antonyms come from? This chapter seeks to address such questions by identifying embryonic antonyms: pairs of words which, corpus evidence suggests, might be developing into the 'opposites' of tomorrow.

Framework productivity

One way of identifying emerging antonyms is to examine productive frameworks associated with new classes of antonymy. Productivity here refers to the 'statistical readiness' (Renouf and Baayen 1996) of lexico-syntactic constructions to incorporate other related terms. In other words, if antonyms occupy certain lexical environments in text, which other words also occupy those environments and could some of those words be seen as new, maturing antonyms?

For example, sixty-two database sentences were assigned to the class of Negated Antonymy. One of the frameworks associated with Negated Antonymy is *X instead of Y*. Therefore, we know that antonyms can occupy X and Y positions in this framework, but we do not know what other relations might be held by words which occupy these positions. It seems likely that, if we return to the corpus and retrieve further occurrences of *X instead of Y*, the missing words in this construction will (sometimes, at least) be contrastive. They may not be familiar antonymous pairs like those included in the database, but context demands that X and Y must be set up in some sort of opposition, even if that opposition is entirely instantial. Such pairs are useful to explore because they reflect all kinds of textual contrasts; they are not just prototypical 'opposites'. And, of course, *X instead of Y* is just one productive framework of antonymy, dozens more of which could be identified. In this way, new antonyms may be identified and, perhaps, diachronically quantified, so that the process by which a pair of words become 'opposites' can be monitored.

Furthermore, once productive frameworks have been identified, it is possible to assess the antonymous profile of a given word by placing it in X-position in one or more framework, then retrieving all corpus occurrences of that word string and examining which terms occupy Y-position. For example, to discover the antonymous profile of, say, *smart*, the corpus could be searched for all instances of *smart instead of* and then a quantified index of all terms which immediately follow this pattern could be developed. This would provide some clues about which words are set up in textual opposition against *smart* most often, though more than one framework would need to be investigated for an accurate impression to emerge.

Of course, this is by no means a foolproof way of identifying antonyms in text. No framework is occupied exclusively by antonyms and potential antonyms; some frameworks are low-frequency, and many words simply generate no recognisable antonymous profile. However, this experiment is useful as a first step towards automatic antonym retrieval and provides some insight into the way in which antonymous pairs become established in language.

The methodology used to test the productivity of antonymous frameworks is based on that pioneered by Renouf (1996) and by Hearst (1998) when testing the productivity of frameworks associated with hyponymy. This strategy involves identifying lexico-syntactic constructions associated with nymic pairs (synonyms, hyponyms, meronyms, etc.), then examining those constructions to discover which other nyms appear in each environment. This methodological approach will now be applied to frameworks associated with new classes of antonymy, most of which were found to favour certain lexical environments in text. For example, the majority of sentences classified in terms of Distinguished Antonymy incorporate the phrase *difference* (or a synonym of *difference*) *between X and Y* where X and Y are antonymous noun phrases. Therefore, the construction *between X and Y* can be regarded as a potentially productive framework because it is possible that other words occupying X and Y-positions will reflect some degree of innate opposition. Indeed, those words which occupy X and Y-positions regularly are likely to reflect some degree of antonymity.

The framework *between X and Y* is an obvious choice because nearly all Distinguished Antonymy sentences conform to this structure. However, Coordinated Antonymy sentences were found to favour not one framework, but several,[1] including *both X and Y*, *either X or Y*, *neither X nor Y*, *whether X or Y*, *how X or Y* and *X and Y alike*. Other new classes which tend to favour certain constructions include Comparative Antonymy (*more X than Y*) and Transitional Antonymy (*from X to Y*). Lower-frequency frameworks which could also be examined in terms of their productivity include *X not Y*, *X rather than Y* and *X as opposed to Y* (Negated Antonymy) and *very X and/or very Y* (Extreme Antonymy). Indeed, the only new classes of antonymy which do not make regular use of productive frameworks are Ancillary Antonymy and Idiomatic Antonymy.

Three of the frameworks mentioned above have been analysed according to

their ability to retrieve contrastive items of a given seed word. Each framework is relatively high-frequency and is associated with a large class of antonymy. Two belong to Coordinated Antonymy and one belongs to Distinguished Antonymy:

both X and Y	(Coordinated Antonymy)
between X and Y	(Distinguished Antonymy)
whether X or Y	(Coordinated Antonymy)

The productivity of these frameworks[2] was tested by placing a word in X-position, extracting all concordances which feature that word string from the entire corpus, then examining which items occupy Y-position most frequently. Three words were placed in X-position for each framework, beginning with an antonym from the database, *good*. For these frameworks to be considered productive in any sense, intuition demanded that they retrieve *bad* and, to a lesser extent, *evil* in Y-position with high frequency. The other two words examined were *natural* and *style*, neither of which have such obvious 'opposites'. This made them ideal candidates for antonymous profiling because they would demonstrate the success with which productive frameworks can be used to generate contrastive output for words which do not yet belong to an established antonymous pair.

Seed word 1: *good*

The first word tested in X-position was *good*, which appears on 181,876 occasions in the corpus. Being among the highest-frequency adjectives in English, *good* was expected to occur in the three frameworks under investigation at a healthy rate. And, being a familiar antonym, it would determine whether these frameworks are productive in terms of retrieving contrastive lexis.

both good and . . .

The word string *both good and* appears in a total of sixty-three corpus sentences. In forty-five of those sixty-three sentences (71.4 per cent), it is followed immediately by *bad*. Thus, sentences such as the following are very common:

95a The procedures recommended by the authors could pass very unfair judgements upon many schools, both **good** and **bad**.
95b Splendidly though he responded, Gascoigne's efforts had the effect of deflecting attention from both **good** and **bad** points about the performances of others.
95c Luck, both **good** and **bad**, tends to even itself out over a period of time.

The above examples are standard Coordinated Antonymy sentences in which

both good and bad serves to exhaust its respective noun head (*schools*, *points* and *luck* respectively). As such, these sentences are typical of those which a productive lexico-syntactic framework would be expected to retrieve. A further four of the sixty-three sentences recorded *evil* appearing immediately after *both good and*. Three of those sentences are listed below:

96a Great possibilities exist for both **good** and **evil**.
96b The forming of human character is a little understood process, but it does, in almost all cases, develop consistencies for both **good** and **evil**.
96c The intense vitality of Gordimer's prose conveys the density and the variousness of human life in a time and place where the old order is rapidly changing, where both **good** and **evil** are possible, where nothing can yet be taken as achieved and yet everything is excitingly there for reshaping and re-inventing: 'Everyone wants their own future arranged around them,' she writes, 'everyone has plans for a structure of laws to contain their ideal existence.'

Once again, it is pleasing to find such contexts being retrieved. In each of the examples above, the antonymous pair function as nouns rather than adjectives, but the effect is the same: to identify a scale and signal inclusiveness. However, this leaves fourteen occurrences of *both good and* which are followed by neither *bad* nor *evil*. These are listed below, together with the noun head they modify:

- both good and flawed (King Hassan's reign)
- both good and friendly (a service)
- both good and great (wines)
- both good and hard (times)
- both good and inimical to the Labour Party (Conservative belief)
- both good and lasting (friends)
- both good and nasty (youths)
- both good and new (a paper)
- both good and non-sexually explicit (a novel)
- both good and not green (God)
- both good and pathetic (years)
- both good and powerful (patriotism)
- both good and true (a story)
- both good and wicked (people)

The concordances above make interesting reading. Some Y-position words are very useful contrast terms for *good*. For example, *flawed* and *pathetic* are valid potential antonyms of *good*, even though neither have become enshrined as an 'opposite' in the same way that *bad* has. As *good* is antonymous with *evil* as well as *bad*, it is pleasing to note that *wicked* and *nasty* have also been retrieved. One

would not intuitively identify *hard* as an antonym of *good*, but this contrast is perfectly valid within its given context – *hard times* are quite the 'opposite' of *good times*. This is a phrasal opposition, whereas most Y-position items reflect single-word contrast, perhaps because the given seed word is also non-phrasal. Another example of a retrieved phrase which initially seems odd, but is actually acceptable in its context is *not green*. Here, *green* means 'environmentally friendly', so a contrast of sorts is being set up regarding ecological issues, although this opposition is more ad hoc than previous examples.

However, the phrase *both X and Y* does not always express an obvious contrast. For example, a story is described as being *both good and true*; one would not want to consider these terms as potential antonyms. In such contexts, it would appear that the framework signals unlikely inclusiveness. Perhaps there is some reason why one would not expect that particular story to be *both good and true*. Alternatively, these words could simply be coupled together without any overt element of contrast being created at all. This appears to be the case with *both good and lasting*. Similarly, no contrast is generated between *good* and either *powerful*, *new*, or *friendly*. These are just further attributes of the thing described as *good*. When Conservative belief is characterised as being *both good and inimical to the Labour Party*, no contrast is generated because being *inimical to the Labour Party* is not at odds with being *good* for Conservative belief. A more borderline example is the novel described as being *both good and non-sexually explicit*. Whether or not these attributes are to be interpreted contrastively is difficult to judge without being privy to the personal preferences of the writer.

Finally, *both good and great* is an interesting example because a distinction is made, but that distinction is not at the usual point on the scale on quality (i.e. between *good* and *bad*). Rather, this context distinguishes between something *good* and something better than *good*. While these terms are instantially contrastive, one could not identify *great* as being a potential antonym of *good*.

between good and . . .

The framework *between good and Y* occurs in 140 corpus sentences, more than twice as many as *both good and Y*. However, the distribution of Y-position output is very different. Of the 140 examples of *between good and Y*, 50 feature *bad* in Y-position, but 78 feature *evil* in Y-position. This is a marked difference from the ratio of *both good and Y*, where ten times as many sentences featured *bad* than *evil*.

This raises the question of why Coordinated Antonymy sentences tend to feature *good*/*bad* (e.g. *both good and bad*) and Distinguished Antonymy sentences tend to feature *good*/*evil* (e.g. *between good and evil*). It would seem that the scale of quality (*good*/*bad*) is used more to signal inclusiveness (usually when referring to people), while distinctions are more commonly made between *good* and *evil* (often in contexts such as a *struggle*, *battle* or *war*). This is confirmed by the fact that Distinguished Antonymy sentences favour nouns over adjectives and Coordinated Antonymy sentences favour adjectives over nouns.[3] Therefore,

when *good* is coordinated with *evil*, the pair are more likely to be nominal (e.g. *both good and evil are possible*), but when *good* is coordinated with *bad*, the pair are more likely to be adjectival (e.g. *many schools, both good and bad*).

Of the 140 sentences which feature *between good and Y*, only 12 feature neither *bad* nor *evil* in Y-position. These contexts are listed below, together with their corresponding noun head, where appropriate:

- between good and excellent (Melbourne's eateries)
- between good and good to soft (the going)
- between good and greed (a struggle within Lewis)
- between good and harmful (foods)
- between good and lousy (comprehensives)
- between good and poor (performance)
- between good and poor (schools)
- between good and really great (wine)
- between good and suspicious (toadstools)
- between good and the best
- between good and very good
- between good and very good

Once again, some of the occurrences at the lower end of the frequency scale are valid contrast terms and others are not. Two contexts show *poor* occupying Y-position in the *between good and Y* framework, when modifying *schools* and *performance*. This is an excellent contrast term, as is *lousy*. Other words are negative, but perhaps more context-specific: *harmful*, when applied to *food*, contrasts with *good*; and the struggle within a boxer between *good* and *greed* is an interesting instantial opposition. In relation to *toadstools*, a contrast between *good* and *suspicious* is acceptable, though it is difficult to imagine this opposition being valid in many other contexts. The distinction between *good* and *good to soft* relates to ground conditions at a horse-racing meeting and is equally context-specific.

Most interesting, perhaps, are the distinctions made between *good* and other, more extreme points on the scale of quality. On two occasions, *very good* is contrasted with *good*, and *excellent*, *really great* and *the best* each appear in opposition on one occasion. These Y-position terms are reminiscent of the *both good and great* word string retrieved earlier. Although one would expect *good* to contrast exclusively with negative items in text, it would seem that many writers choose to exploit its latent contrast with 'super-positive' terms instead.

whether good or . . .

Of the three lexico-syntactic frameworks analysed, *whether X or Y* is the least common. In the corpus, only eight sentences feature the word string *whether good or Y*. In seven of those sentences, *bad* fills the Y-position; in the eighth, *evil* fills the Y-position. This bias toward *bad* is not unpredictable given that

Coordinated Antonymy sentences have already been shown to favour *bad* over *evil*. It is interesting to note that no other words fill Y-position in the framework. This may suggest that non-standard oppositions avoid this construction, although, given that it only occurs in eight sentences out of 12.8 million, such a conclusion may be prematurely drawn.

Antonymous profile of good

This analysis of *good* has confirmed that it is definitely possible to retrieve contrast words from the corpus using productive lexico-syntactic frameworks. Collectively, the three frameworks examined occur on 236 occasions in the corpus. On 114 of those occasions (48.3 per cent), the given word string is followed by *bad*. This is compatible with our intuitions – one could predict that *bad* would be set up in opposition against *good* most commonly. Indeed, one could also predict that *evil* would be runner-up; *evil* occupies the Y-position in frameworks analysed on eighty-eight occasions (37.3 per cent). However, the purpose of this experiment was not to prove that *bad* and *evil* are antonymous with *good*; rather, it was to show that the three frameworks identified are fertile enough to be deemed productive. This seems indisputable.

Interestingly, *bad* and *evil* were not the only terms retrieved from the corpus using this method. A further thirty-four sentences (14.4 per cent) yielded a word or phrase other than an established antonym. Some of these were interesting potential 'opposites' of *good*; others were more context-dependent distinctions. Useful contrast words included *wicked*, *nasty*, *pathetic*, *poor* (twice), *lousy* and *flawed*, but among those words which would not be intuitively recognised as antonyms of *good* were a group of terms at the extreme of the quality scale (such as *great*, *really great* and *excellent*), as well as words which lack any obvious contrast (such as *true*, *friendly* and *lasting*). However, in total, over 90 per cent of items extracted using the three lexico-syntactic frameworks yielded potential contrast terms for the X-position word. This demonstrates that all three frameworks are productive and suggests that they may also prove successful in investigating the antonymous profile of words which have no established 'opposite' in English.

Seed word 2: *natural*

The next word placed in X-position was *natural*, which occurs on 22,920 occasions in the corpus. If pressed to give an 'opposite', most speakers would probably opt for *unnatural*, but this is by no means the only potential antonym of *natural*.

both natural and . . .

The output generated by a search for *both natural and* comprises thirty-three concordances. Each of these is recorded below, together with the adjective which follows this word string and the noun head to which the words refer:

- both natural and accurate (their response to the camera)
- both natural and artificial (everything that exists)
- both natural and artificial (light)
- both natural and artificial (light)
- both natural and artificial (lighting)
- both natural and artificial (the essence of man)
- both natural and assisted (fertility)
- both natural and beneficial (high altitude)
- both natural and coloured (light)
- both natural and heraldic (devices)
- both natural and human (perturbations)
- both natural and inevitable (disaster)
- both natural and inevitable (process)
- both natural and lucid (her acting)
- both natural and man-made (beauty)
- both natural and man-made (beauty)
- both natural and man-made (components)
- both natural and man-made (disasters)
- both natural and man-made (facilities)
- both natural and man-made (polymers)
- both natural and market (forces)
- both natural and prudent (paying debts)
- both natural and safe (white sugar)
- both natural and sensible (idea)
- both natural and social (sciences)
- both natural and social (sciences)
- both natural and spiritual (creatures)
- both natural and superb (a history of vodka)
- both natural and synthetic (fibres)
- both natural and taboo (a child's sexuality)
- both natural and technical (the effect)
- both natural and violent (causes)
- both natural and vital (USA action)

It can be seen that *both natural and* occurs less frequently in text than *both good and* (thirty-three hits of the former compared to sixty-three hits of the latter), and that the Y-position output is more diverse. However, this is not to say that no patterns emerge in the concordances above: of the thirty-three occurrences of this lexico-syntactic framework, six are followed by *man-made* and five are followed by *artificial*. Both of these terms make excellent contrast words for *natural*.

Retrieved on two occasions each are *inevitable* and *social*. Neither initially appeal as being good 'opposites' of *natural*, but, within its given context, *social* is entirely valid. This is because of the common academic distinction between natural sciences and social sciences. *Inevitable* works less well as a contrast word of *natural*, suggesting that the *both X and Y* framework functions here

only in terms of inclusiveness. Indeed, many of the words retrieved in Y-position on one occasion only are also non-contrastive. However, interesting and valid oppositions of *natural* include *market* (in terms of forces), *synthetic* (in terms of fibres), *violent* (in terms of death) and *assisted* (in terms of fertility).

between natural and . . .

- between natural and artificial (ozone)
- between natural and artificial (worlds)
- between natural and artificial (worlds)
- between natural and created (forms)
- between natural and cultivated (areas)
- between natural and juridical (persons)
- between natural and man-made (assets)
- between natural and metal (packaging)
- between natural and moral (evil)
- between natural and supernatural

Ten sentences feature the phrase *between natural and Y*, significantly less than the 140 sentences which featured *between good and Y*. All ten concordances are recorded above and it can be seen that the only word to occupy Y-position more than once is *artificial* (once when applied to *ozone* and twice when applied to *worlds*). This confirms that *artificial* is commonly set up in opposition against *natural* in text. It is interesting to note that *between natural and man-made* also appears. This suggests that *man-made* shares a similarly strong contrastive profile with *natural*. Significantly, it also confirms that similar words occupy Y-position in the frameworks *both natural and Y* and *between natural and Y*.

All of the six words that occur only once in Y-position reflect contrast to a lesser degree, with *supernatural* perhaps being the most interesting, especially because it ties in with *spiritual*, which was picked up by *both natural and Y*. *Created* and *cultivated* are both relatively synonymous with *man-made*, and *metal* is another example of a potentially valid, but unpredictable contrast (within the field of packaging, *metal* and *natural* may be a more familiar opposition than the lay-person realises).

whether natural or . . .

- whether natural or artificial (hormones)
- whether natural or electric (light)
- whether natural or imposed (punishment)
- whether natural or man-made (environment)
- whether natural or man-made (beauty)
- whether natural or otherwise (phenomena)

- whether natural or step (parents)
- whether natural or through external intervention (chemical changes)

In terms of quantity, *whether X or Y* is the least frequent of the three frameworks examined. However, the quality of the contrast words generated is not compromised, as the eight concordances above illustrate. Pleasingly, both *artificial* and *man-made* arise in Y-position again: the former on one occasion (noun head: *hormones*); the latter on two occasions (noun heads: *environment* and *beauty*). This means that all three lexico-syntactic frameworks have successfully retrieved both of these words, testimony to the effectiveness of the methodological approach used.

One-off contrast terms again include valid and useful examples. For example, *step* is not the kind of word one would intuitively identify as a potential antonym of *natural*. However, within the given context of parentage, this contrast is not only legitimate but very interesting. It is also reassuring to note the appearance of *otherwise* in Y-position. Though this term is not a valid 'opposite' of *natural* in itself, *otherwise* effectively functions as a proform for unspecified contrast words in text.[4] Finally, the description of chemical changes (*whether natural or through external intervention*) is significant because it illustrates that contrast terms are not restricted to single words. In this context, *through external intervention* is as valid an opposition as *artificial* or *man-made*.

Antonymous profile of natural

Collectively, the three frameworks examined feature *natural* in X-position in a total of fifty-one sentences. In nine of those sentences, *artificial* occupies Y-position and, in a further nine of those sentences, *man-made* occupies Y-position. This strongly suggests that those two words are the primary textual contrast terms of *natural*. Indeed, over one-third of all frameworks examined feature either *artificial* or *man-made*. No other term is retrieved on more than two occasions from the corpus. As with *good*, low-frequency items tend to be a mixed bunch. However, a number of interesting potential contrast words for *natural* were found, including *synthetic*, *created* and *imposed*.

This output is particularly interesting if analysed in the light of the range of antonyms which lexicographers have paired intuitively with *natural*. For example, *Webster's Dictionary of Synonyms* (1951) lists three antonyms: *artificial*, *adventitious* and *unnatural*. The inclusion of the first-mentioned term is supported by this experiment, but *adventitious* does not occupy Y-position at all. This is not surprising given that the word occurs only seven times in the entire corpus (or about once per 40 million words in text). More surprising is the non-appearance of *unnatural* in textual opposition with *natural.* This reveals an interesting fact about *natural*: namely, that it prefers to contrast with lexical antonyms rather than its morphological antonym. Indeed, if we broaden the scope of this experiment to look at all *X and/or Y* contexts, we find that *natural and/or unnatural* occurs thirteen times in the corpus, but *natural*

and/or artificial appears twenty-five times and *natural and/or man-made* appears forty-three times. It would be interesting to discover whether other adjectives follow this trend of favouring non-morphological contrast words even when a morphological antonym is available.

Collins Cobuild Dictionary (1987)[5] cites six antonyms of *natural*, beginning with *unnatural*. The other contrast words suggested are *surprising* (not retrieved in text), *contrived* (not retrieved), *artificial* (retrieved nine times), *man-made* (retrieved nine times), and *processed* (not retrieved). *Chambers Dictionary of Synonyms and Antonyms* (1989) suggests *unnatural*, *artificial*, *man-made*, *affected* and *contrived*.

Therefore, some correlation emerges between intuitively identified antonyms and antonyms identified by productive lexico-syntactic frameworks: all three dictionaries cite *artificial* as a likely antonym and only the oldest of the three fails to cite *man-made*. However, I would suggest that other recommended antonyms (*adventitious*, *processed* and even *unnatural*) are not placed in textual opposition against *natural* as often as may have been anticipated. Moreover, it could be argued that such words are less valid contrast terms of *natural* than *synthetic*, *supernatural*, *assisted* and other Y-position words which have not been cited by lexicographers, but which were retrieved in this experiment.

Seed word 3: *style*

The final word to be placed in X-position was *style*, which occurs on 34,029 occasions in the corpus and could not be described as a recognised antonym of any other word. Being a noun rather than an adjective, *style* would test whether the frameworks identified are robust enough to house words belonging to more than one word class.

both style and . . .

- both style and a demonstration of reaching speed
- both style and achievement
- both style and commercial space
- both style and content (× 4)
- both style and date
- both style and emotion
- both style and fashion
- both style and feeling
- both style and heart
- both style and history
- both style and performance
- both style and personality
- both style and personnel
- both style and policy

- both style and prices
- both style and qualifications
- both style and reputation
- both style and standards
- both style and substance (× 5)

The framework *both style and Y* occurs in twenty-six corpus sentences and, as the concordances above show, the words which fill Y-position are diverse. However, two terms are set up in opposition against *style* more frequently than any others: *substance* appears in Y-position five times and *content* appears in Y-position four times. These contexts reflect a trend for *style* to be seen as meaningless or superficial, and licenses its opposition with more 'weighty' terms. Other words which reflect this trend but only occur once in Y-position include *performance*, *policy*, *achievement* and *standards*. From examining the output above, one might infer that *style* is developing a pejorative sense in the language, an impression confirmed by the occurrence of *emotion*, *feeling* and *heart* in Y-position. This sub-set of 'passion' terms suggest that *style* is sometimes felt to be incompatible with strong emotions.

between style and . . .

- between style and content (× 4)
- between style and disorder
- between style and grape
- between style and political ideology
- between style and quality (× 2)
- between style and subject
- between style and substance (× 5)

As recorded above, the word string *between style and* appears fifteen times in the corpus: on four occasions it is followed by *content*; on five occasions it is followed by *substance*; and on six occasions it is followed by other terms, including *quality*, which conforms to the underlying trend for *style* to be treated negatively in text. However, a reminder that these frameworks are not exclusively inhabited by contrast words is provided by *grape*. The sentence from which this word string is taken actually explores the relationship between the style of a given wine and the nature of the grape used in its production. However, *grape* could hardly be seen as a valid and useful 'opposite' of *style*, merely an unlikely instantial collocation.

whether style or . . .

This lexico-syntactic framework was not found in the corpus at all, probably because *style* functions most commonly as a noun and nouns do not lend themselves readily to this construction. This raises the issue of whether it

remains a valid productive framework for antonyms. However, just because a framework is not receptive to all word classes, it does not follow that the framework is not useful in yielding interesting contrast words. The criterion is quality of contrast, not quantity.

Antonymous profile of style

The antonymous profile of *style* is intriguing. The framework *both style and Y* occurs twenty-one times in the corpus; the framework *between style and Y* occurs fifteen times; and the framework *whether style or Y* does not occur at all. Most commonly retrieved in Y-position is *substance* (ten hits; 24.4 per cent). In second place is *content* (eight hits; 19.5 per cent). Between them, these two words are retrieved in Y-position in 43.9 per cent of frameworks. This suggests that *style* tends to be set up in opposition against things with greater intellectual depth and is regarded as being superficial in many respects. This negative use of *style* is not acknowledged by many modern dictionaries, perhaps because it only becomes apparent when one examines its antonymous profile in text. Some of the low-frequency words retrieved in Y-position also evidence *style* being treated as a negative attribute. On two occasions, *quality* is contrasted with *style* and words such as *policy* and *political ideology* also seem to be considered incompatible with *style*. It is interesting to note that *style* is never seen to contrast with concepts such as *inelegance* or *tastelessness*, as *Chambers Dictionary of Synonyms and Antonyms* (1989) suggests it might.

Initially, the opposition between *style* and *substance* (or *content*) tended to be used most often in the context of New Labour, and the current British prime minister, Tony Blair. However, more recently, this opposition has begun to move away from the political spectrum and be used in any context[6] in which the appearance of something or somebody seems at odds with their performance or ability. This suggests that *style* and *substance* are receiving wider exposure in language and, perhaps, edging higher up the scale of antonymity. As the corpus consists of newspaper stories published between 1988 and 1996, the fact that *style* and *substance* seem to be operating more extensively and more flexibly in 2002 indicates that word pairs may become antonymised very rapidly.

Is antonymous profiling possible?

The process by which 'opposites' are created is complex, but it is reasonable to speculate that in order for a pair of words to become enshrined as antonyms in any language, they must first receive a significant amount of exposure. This exposure is likely to be in contexts which are more often associated with established pairs of antonyms. Therefore, this chapter has sought to determine whether any of the lexico-syntactic frameworks associated with the various new classes of antonymy could be used to assess the antonymous profile of a given seed word in journalistic text. A number of frameworks could have been

chosen, but this study has investigated the productivity of three common constructions: *both X and Y*, *between X and Y* and *whether X or Y*.

The first word to be placed in X-position was *good*, which, as one would expect, retrieved *bad* and, to a lesser extent, *evil* in Y-position. This output augured well for the productivity of the frameworks selected, and justified the examination of other seed words. An analysis of *natural* showed that two words were set up in opposition more often than any others. Indeed, between them, *artificial* and *man-made* occurred in Y-position in over a third of all contexts retrieved. Finally, an analysis of *style* also revealed a pair of terms competing with one another to become the favoured textual opposition: *substance* was retrieved on ten occasions and *content* on eight occasions.

This may be seen as initial evidence that the automatic retrieval of antonyms in text is possible using some of the lexico-syntactic frameworks associated with new classes of antonymy. Further research would be necessary to identify which strategies are most efficient: certain word classes could prove to be more inclined to certain frameworks; and certain new classes could prove more productive than others. However, the results presented here are sufficient to demonstrate the productivity of three frameworks in text and much of the output generated (both high- and low-frequency) would be of interest to dictionary-makers who currently appear to identify antonyms on an exclusively intuitive basis.

However, one could argue that this research does not succeed in shedding light on the true origin of antonyms, but, rather, is simply a method of identifying antonyms once they have surfaced in language. In other words, we know that, say, *style* and *substance* seem to be developing antonymous status, but we do not know why. Presumably, this has something to do with demand being created for a pair of words to express such an opposition, a demand which, in this case, could be symptomatic of the perceived 'dumbing down' of British politics. Indeed, to return to arguably the most dramatic antonymous formation of recent times, *gay* and *straight*, the increasing awareness of and attention given to homosexuality surely contributed to the establishment of these words as a new 'opposites'. However, it is reasonable to assume that this antonymity was affirmed, enshrined and disseminated by the co-occurrence of *gay* and *straight* in lexical environments similar to those examined in this chapter.

11 Antonymy
Past, present and future

By now, the success with which corpora can be used to examine 'opposites' should be apparent. Primarily, various functions of antonymy within a 3,000-sentence database have been identified, quantified and used as the basis for a new system of classification. However, this database has also enabled more subtle aspects of antonymy to be explored. For example, the fact that most antonymous pairs favour a given sequence in text has been established and discussed, as has the influence of word class on the function of antonymy, and the question of whether non-gradable antonyms operate in the same way as gradable antonyms. This research can be seen as a complement to pre-corpus investigations of antonymy and as a precursor to further data-based analyses. The purpose of this chapter is threefold: to look back at past investigations of antonymy so that the relationship between traditional classes and new classes can be assessed; to compare this research with other corpus-based studies of antonymy so that the present position about antonymy can be consolidated; and, finally, to make explicit the limitations of this study so that future research may focus on those aspects of antonymy not adequately explored here.

Antonymy past

Antonymy has invariably been packaged together with a batch of other sense relations for research and teaching purposes. However, thinking about antonymy alongside synonymy, hyponymy, meronymy, etc. obscures the fact that these relations are all diverse in nature. Unlike other sense relations, 'opposites' are a key element of human communication, becoming entrenched in the psyche from a very early age and playing an important part in our daily lives. Antonymy is not just another member of the nymic family, as traditional semantic theory sometimes implies.

The reason why antonymy infiltrates our consciousness so deeply is not easy to pinpoint. Antonyms belong to a two-member system (as opposed to synonymy, where *sofa*, *settee* and *couch* could all be seen as similar in meaning) and each member is faithful to the other (as opposed to hyponymy, where *dog* is one of many hyponyms of *animal* and can itself act as a superordinate term). Though antonymous profiling shows that a given word may be set up in

opposition against a number of contrastive items in text, antonyms generally have only one established partner[1] and this may help the pair to become fixed in our mental lexicon. Two-member systems seem able to form closer associations than multi-member systems, even when non-antonymous, as illustrated by irreversible binomials such as *now and then*, *rest and recuperation*, *salt and vinegar*, etc.

However, there is more to antonymy than collocation. Lyons quotes the German semanticist Trier as stating that 'every word that is pronounced calls forth its opposite' (1977: 270). This may sound hyperbolic, but it begins to explain the unique relationship holding between antonyms. Language reflects the urge we have to dissect, an urge which may be related to the fact that we are all either *female* or *male*. That dichotomy is natural, but other dichotomies are more artificial. For example, people are still labelled as being *black* or *white*, despite the multitude of skin colours which abound (none of which, ironically, are truly black or white) and the frequent unhelpfulness of the distinction. It is as though, as Trier suggests, we conceptualise words, not just in terms of what they mean, but in terms of what they do not mean. Therefore, knowing the antonym of a given word or concept enhances our understanding of that word or concept. This may explain why terms are sometimes coined with no obvious function other than to act as an antonym. For example, *able-bodied* is increasingly used as an antonym of *handicapped* and neologisms such as *undelete* and *unsubscribe* are becoming familiar Internet terms. Humans seem prone to organise the world in terms of oppositions; antonymy is simply a linguistic reflection of this. Lyons argues that 'binary opposition is one of the most important principles governing the structure of language' (1977: 271) and corpus-based research provides no grounds to dispute this statement. So, though it is convenient to handle antonymy under the umbrella term of 'sense relations', it should be acknowledged that this superordinate may be deceptive. Unlike synonymy (which does not truly exist) and hyponymy and meronymy (essentially relations imposed on the language by analysts but unbeknown to the user), antonyms surround us and form an integral part of the human communication system. That is why English, like most languages,[2] has a non-technical word for antonyms: 'opposites'.

However, even though the context in which antonymy has traditionally been placed is inappropriate in many respects, there is no doubt that 'opposites' have received considerable linguistic attention in the latter half of the twentieth century, usually being classified according to their inherent semantic properties. For example, amongst other distinctions, gradable antonyms (*cold*/*hot*, *long*/*short*, *hard*/*soft*, etc.) have been differentiated from non-gradable antonyms (*alive*/*dead*, *female*/*male*, *illegal*/*legal*, etc.). This distinction usually comes with a disclaimer: most non-gradable antonyms are, in fact, commonly graded in text. However, the distinction may also be challenged from another perspective, namely the tendency of non-gradable pairs to function in a similar way to gradable pairs. Admittedly, some correlation arises between the textual function of a pair of antonyms and their gradable

status, but this correlation is relatively weak. All eight major new classes of antonymy are served by both gradable and non-gradable pairs. This supports the view that antonym usage is largely unrelated to the innate semantic properties of individual pairs.

To give another example of traditional theory being amenable to a more text-driven approach, sentences which follow a *how X or Y* framework challenge the widely held view that an antonymous pair has a marked and unmarked term. Accepted semantic wisdom holds that one antonym can be used in an unmarked fashion ('how high is that building?') while its partner remains marked ('how short is that building?'). However, this widespread belief is questioned by sentences which use the construction *how X or Y* (e.g. 'how well or badly is the NHS doing?'; 'how easy or difficult is the assignment?'). This framework, though comparatively rare, suggests that writers do not always feel that the unmarked antonym is truly unmarked.

Furthermore, the focus of many early analyses of antonymy was on adjectival pairs,[3] with nouns and verbs being used only to exemplify relational oppositions, such as reciprocality, and adverbial pairs being largely ignored. This is counter-intuitive (if *rich* and *poor* are antonymous, why aren't *wealth* and *poverty?*) and misleading (*wealth* and *poverty* actually serve many of the same textual functions as *rich* and *poor*). The two most common classes were found to be Coordinated and Ancillary Antonymy, regardless of whether the pairs examined were adjectives, nouns, verbs or adverbs. Indeed, most frameworks associated with antonymy are able to house antonyms belonging to any of these word classes. Though some exceptions were found (a minority of individual antonymous pairs display irregular textual behaviour), this indicates that semantics takes precedence over grammar; that the antonymity of these pairs dictates their usage, not their grammatical status. To focus only on adjectives (and, remember, the much-vaunted gradable/non-gradable distinction only really applies to adjectives) is to leave hundreds of valid antonymous pairs unexplored.

Those intuitively developed categories of antonymy have criteria which can now, thanks to corpus data, be put to the test. For example, when defining the category of 'complementaries', Cruse states that such pairs 'can be diagnosed by the anomalous nature of a sentence denying both terms' (1986: 198). He then inserts various non-gradable pairs into the framework *neither X nor Y* and claims that such contexts are infeasible. However, the corpus yields the following counter-examples, which suggest that this criterion is far from watertight.

97a In a Birkbeck tutorial, a second-term group was wrestling with a theory that statements can be neither **true** nor **false**.
97b The door is neither **open** nor **closed**.
97c It is a bit harsh to write this unpretentious whodunnit off as 'dead theatre', for really it is neither **alive** nor **dead**, but in limbo.

98a Jobs will just be jobs, neither **male** nor **female**.

98b Many of the Free churches now ordain women, following the scriptural tenet that 'in Christ there is neither **male** nor **female**'.

98c To walk in and say 'I'm sorry I'm late' was neither **male** nor **female**: it was just grown-up.

The six sentences above all show a non-gradable pair occurring in a *neither X nor Y* framework, something which Cruse implies is logically unsound. In the first triplet of sentences, these pairs are *false*/*true*, *closed*/*open* and *alive*/*dead*. In the second triplet of sentences, the pair is *female*/*male*, which occupies this framework quite frequently. These examples show that intuitive criteria are not ideal as a means to distinguish between pairs. Contexts which are unacceptable in one mental lexicon may not be so readily rejected by all writers, and corpus data once again raises questions about the methodological validity of classifying antonymy from an introspective perspective.

Of course, in isolation, these arguments hardly render useless the entire body of literature about antonymy. Many analyses which rely on little more than personal intuition remain detailed and insightful. Indeed, it is remarkable how many intuitive statements about antonymy are affirmed by textual evidence. However, just as pre-corpus dictionaries were inevitably improved once it became possible to identify how individual words actually operate in text, so too are logical classifications of antonymy receptive to data-based reappraisal. A corpus-based study of antonymy does not displace traditional classes of antonymy, it runs parallel to earlier research and provides an alternative view of the phenomenon.

Antonymy present

This investigation of antonymy was based on an analysis of 3,000 sentences, each featuring both members of an antonymous pair, randomly extracted from a 280-million-word corpus of journalistic text. When this database was manually classified according to the textual function of each word pair, it was found that 96.5 per cent of database sentences could be attributed to one of eight classes. Two of those classes dominate the distribution: Ancillary Antonymy accounts for 1,162 sentences and Coordinated Antonymy accounts for 1,151 sentences. Collectively, these two functions account for antonymous pairs in 77.1 per cent of contexts. This distribution contradicts Fellbaum's statement that 'the contexts of [antonymous] co-occurrences are formally unstructured and the syntactic patterns vary widely' (1995: 294), although this claim was based upon evidence from a very small corpus. Table 11.1 is a reminder of all eight classes, together with an example for each taken from the *new*/*old* component of the database (*new*/*old* being the most commonly co-occurring antonymous pair in newspaper text).

In general, one could say that writers choose to exploit our familiarity with antonymy. More than other sense relations, 'opposites' are instantly recognised by language users and writers 'cash in' on the status held by an

Table 11.1 Summary of new classes

Ancillary Antonymy	(38.7%)	. . . the present charts are either filled with **old** *songs* or **new** *rubbish*.
Coordinated Antonymy	(38.4%)	Paganism, **old** and **new**, sees a world that is disorderly and irrational.
Comparative Antonymy	(6.8%)	. . . the **new** Webb is considerably more wonderful than the **old**.
Distinguished Antonymy	(5.4%)	The main difference between the **new** contracts and the **old** is that . . .
Transitional Antonymy	(3.0%)	. . . like a dramatisation of the switch from the **old** notation to the **new**.
Negated Antonymy	(2.1%)	. . . the fact of the matter is that these ideas aren't **new**, they're **old**.
Extreme Antonymy	(1.3%)	. . . stands for both the very **new** in politics, and for the very **old** . . .
Idiomatic Antonymy	(0.8%)	. . . he has had no trouble teaching this **old** dog **new** tricks.

antonymous pair when they use it to act as a lexical signal (as in Ancillary Antonymy sentences) or to act as a signal of exhaustiveness or inclusiveness (as in Coordinated Antonymy sentences). Antonymy occasionally occurs in other, perhaps more predictable ways, such as to mark a distinction, comparison or transition, but these contexts are relatively rare.

Corpora have been used in previous investigations of antonymy and, as one would expect, functions of antonymy similar to those identified here have been noted. The advantage that this research has over earlier data-based investigations is that it classifies pairs according to usage and provides new terminology with which to talk about antonymy. It also benefits from being based on a corpus much larger than those used previously. Furthermore, it is now possible to compare these new classes with those functions of antonymy identified elsewhere, with a view to establishing whether different corpora provide a consistent picture of how antonyms operate. A reasonable question to ask of this classificatory system is: are new classes of antonymy sufficiently robust to accommodate contexts retrieved from other corpora by other analysts as part of other investigations? This question will be answered with reference to research undertaken by Justeson and Katz (1991), Fellbaum (1995) and Mettinger (1994).

Justeson and Katz (1991) cite a total of twenty-one contexts[4] which feature an established antonymous pair. In twelve of those contexts, the antonymous pair are functioning in a coordinated environment. Three such examples, retrieved by Justeson and Katz, are recorded below.

99a The welfare of citizens, **old** and **young**, is the responsibility of the community.

99b. The religion, in fact, is an expression of the unity of the group, **small** or **large**.

99c. To us and every nation of the Free World, **rich** or **poor**, these questions are necessary today as never before if we are to march together to a greater security, prosperity and peace.

Though taken from a different corpus by different researchers, the contexts above are clearly akin to those database sentences assigned to the class of Coordinated Antonymy. Justeson and Katz prefer to describe this use of antonymy by stating that 'when antonyms co-occur, it is usually by substituting for one another in phrases that are otherwise repeated word-for-word within the sentence' (1991: 143). The problem with talking about substitution is that antonyms are not entirely substitutable for one another in text. As can be seen in Table 7.1, most pairs of antonyms tend to follow a given sequence within the sentence and, although substituting antonym for antonym would not result in ungrammaticality, this sequence is not usually reversed. Describing these contexts as 'coordinated' seems preferable to describing them in terms of substitution.

Among the other nine antonymous contexts cited by Justeson and Katz are examples of Distinguished Antonymy, Transitional Antonymy and Ancillary Antonymy. The ancillary function of antonymy can be seen at work in the sentences below, all of which were cited by Justeson and Katz.

100a It certainly looked as if *their own Congregationalists* were **wrong** and *the Baptists* **right**.

100b Her *subject* was **large** – *a copy of the Last Supper* – and her *canvas* **small** – *the head of a tiny screw*.

100c It *assumes that things are as they seem* when they seem **best**, and when they seem **worst** it *overlooks them*.

This triplet of sentences exemplify Ancillary Antonymy well. In sentence 100a, *right* and *wrong* signal that *their own Congregationalists* should be contrasted with *the Baptists*. Sentence 100b is more complex because it features two B-pairs: *large* and *small* help us to see *her subject* (a copy of the Last Supper) and *her canvas* (the head of a tiny screw) in opposition. Finally, *best* and *worst* mark a contrast between clauses in sentence 100c. Justeson and Katz analyse these contexts in terms of antonym substitution and do not differentiate between Ancillary and Coordinated Antonymy. However, it seems evident that the function of antonyms in sentences 99a–c differs from the function of antonyms in sentences 100a–c.

Examples of Ancillary Antonymy are also cited by Fellbaum (1995) among the thirty-five sentences from the Brown corpus which she records when examining co-occurrence and antonymy. Indeed, even though relatively few antonymous contexts are analysed, five of the eight new classes outlined here are represented. Fellbaum does not formulate any system of classification, but

uses these sentences to demonstrate that antonyms have 'different syntactic environments and are not substitutable for one another' (1995: 821). Below are five contexts cited by Fellbaum (1995) and the new classes of antonymy to which they would now be attributed:

Ancillary:	We **opened** on *Friday* and **closed** *the following Monday*.
Coordinated:	. . . one more reason she didn't look forward to Cathy's visit, **short** or **long**.
Distinguished:	. . . the outcome may mean the difference between **life** or **death**.
Transitive:	. . . the beatification of America's first saint … was a fiasco from **start** to **finish**.
Negated:	. . . the bombs dropped on Japan were to **end** a war, not **start** one.

Mettinger (1994) cites many more antonymous contexts than either Justeson and Katz (1991) or Fellbaum (1995), but his examples remain compatible with the classes proposed here. Below are eight contexts taken from Mettinger's corpus of modern fiction, accompanied by the new classes to which they would be assigned:

Ancillary:	Where *Miss Milray* was **hard**, *she* was **soft** . . . (1994: 63).
Coordinated:	. . . girls **fat** and **thin**, **short** and **tall**, **clean** and **dirty** . . . (1994: 35).
Comparative:	But it's the **usual** rather than the **unusual** thing to happen (1994: 39).
Distinguished:	. . . everyone knew the difference between **right** and **wrong** (1994: 44).
Transitive:	Someone has turned the **comedy** into a **tragedy** (1994: 37).
Negated:	It's no longer **attack** but **defence** (1994: 37).
Extreme:	Either this crime was very **simple** . . . or extremely **complex**. (1994: 50).
Idiomatic:	He looked Poirot **up** and **down** . . . (1994: 48).

All eight major classes of antonymy are exemplified above, but Mettinger prefers to describe these contexts in terms of 'frames' such as *X and Y*, *X not Y*, *X rather than Y* and *from X to Y*. Though these frames approximate to the new classes of Coordinated, Negated, Comparative and Transitional Antonymy respectively, Mettinger provides no frequency information and, like Justeson and Katz (1991) and Fellbaum (1995), does not isolate contexts in which antonyms function in an ancillary fashion. Nevertheless, it is pleasing to note that the classes presented here are able to account for antonymous usage retrieved from other corpora and cited by other analysts. This suggests that a consistent description of the textual function of antonymy has emerged.

Antonymy future

Before considering as yet uncharted aspects of antonymy, it should be acknowledged that this research, like most research, has its imperfections. Drawing attention to these weaknesses is one way of identifying areas in which future investigations may be able to improve and elaborate this research. For example, the corpus used in this study consists of 280 million words of newspaper text, spanning the period from October 1988 to December 1996. By current standards, this corpus is difficult to fault in terms of size, but one could argue that journalistic text is not representative of language in general. Certainly, the role served by antonymy in speech is not considered here, which is a matter of some concern because spoken language has primacy over written language. On the other hand, newspapers are a good reflection of current usage and no corpus can ever guarantee total representativity.

The database derived from this corpus could be criticised from a number of angles. The fifty-six antonymous pairs selected for study were chosen largely according to personal intuition and the number of examples retrieved for each pair was also determined manually. In some respects, this makes a mockery of objectivity. However, such criticism should be tempered by the fact that studying antonymy is more problematic than studying, say, modal auxiliary verbs, which are finite in number and can be sampled accordingly. Antonymous pairs are impossible to quantify because they can only be defined by consensus. What might be identified as a good pair of 'opposites' by one speaker may be dismissed by the next, so no definitive list of antonyms in English will ever exist. However, this study, like all others, required a starting point. Therefore, it was decided that the sample should be created intuitively, but according to certain criteria regarding the proportion of pairs chosen which are non-gradable and gradable; morphological and lexical; adjectival and non-adjectival. Inevitably, this leaves the study vulnerable to criticism that it is based on an unrepresentative sample of antonymous pairs, but any corpus-based analysis of antonymy needs to be based on something and this sample is among the largest and most diverse used in any investigation.

One might also argue that, though using corpus data is more scientific than not using corpus data, the classes of antonymy presented here are still subject to the interpretative whims of one analyst. For example, despite looking at real data, neither Justeson and Katz (1991) nor Mettinger (1994) identify the function which I refer to as Ancillary Antonymy. On a smaller scale, individual sentences can cause problems regarding which class of antonymy they should rightfully be assigned to. The database examples below illustrate some of the more subjective choices which needed to be made.

101a However, critics say the league tables encourage schools to concentrate on getting the maximum number of pupils through five GCSEs at top grades and are increasing polarisation, with a widening gap between pupils who do **well** and those who do **badly**.

101b It was **sad** but **happy** that Ashley Lawrence should have been among dancers when he died.

101c For years her book has seemed a **minor** classic; now perhaps it will be a **major** one.

Sentence 101a exemplifies an interesting classificatory dilemma. Here, *pupils who do well* are contrasted with *those {pupils} who do badly*. This is preceded by reference to a *widening gap*, which would usually be seen as a signal of Distinguished Antonymy. However, one could also argue that this antonymy is ancillary because *well* and *badly* signal an opposition between two different groups of pupils, each linguistically represented by *who*. When classifying this sentence, I adopted the latter perspective, reasoning that, in the real world, this sentence alludes to two distinct referents and that *well* and *badly* mark a contrast between them. However, one could counter-argue that this example features a framework which is typical of Distinguished Antonymy: *gap between X and Y* where X and Y are antonymous noun phrases (*pupils who do well* and *those {pupils} who do badly*). Neither interpretation is unsound; one can only strive to be consistent when classifying awkward sentences.

At first glance, sentence 101b seems to belong to the class of Coordinated Antonymy with *but* being used in an unusual fashion[5] to signal that something is, inclusively, both *sad* and *happy*. However, this example has actually been classified in terms of Ancillary Antonymy, even though the B-pair is not easily identified or italicised. My view is that *sad* seems related to the fact that Ashley Lawrence died and *happy* seems related to the fact that he was among dancers at the time. However, not every analyst would have interpreted this sentence in this way.

Finally, sentence 101c is on the border of Ancillary Antonymy and Transitional Antonymy. This example has been classified in terms of the latter because of the temporal movement made by the book in question. However, one could claim that an element of ancillary contrast is present here; that *minor* is related to the period of time associated with *for years* and that *major* is associated with *now*. Once again, both interpretations are justifiable, but subjective decisions are unavoidable when dealing with the small minority of sentences which do not fit snugly into any one class.

Indeed, occasionally, ambiguity may arise regarding the class of antonymy to which a whole group of sentences rightfully belong. For example, twenty-three database contexts record antonyms occurring in a *neither X nor Y* framework. These sentences have been attributed to the class of Coordinated Antonymy, but one could argue that they are, in fact, examples of Extreme Antonymy. Similarly, antonyms which occur in an *X rather than Y* framework have been assigned to Comparative Antonymy, not Negated Antonymy; and the entire sub-class of Equal Comparison is regarded as a branch of Comparative Antonymy, not Coordinated Antonymy.[6]

Therefore, although corpora allow for antonymy to be approached in a more objective light, data is still subject to individual interpretation. This is not a

methodological flaw (in some respects it is a methodological necessity), but it would be interesting to discover whether other analysts, given identical data, would distribute sentences among similar classes (and in similar proportions) to those recorded here. Of course, most corpus-based research could be criticised on similar grounds. Perhaps the best defence is transparency: this analysis of the database is open to scrutiny and the classificatory choices made can be evaluated and, if necessary, challenged; pre-corpus categories of antonymy are based on individual mental lexicons, not all of which are available for subsequent interrogation.

So which aspects of antonymy remain unexplored? Well, this database does not include all kinds of antonymous pairs. For example, adjectives are only sampled in their base form (*heavy*/*light*), never in their comparative (*heavier*/*lighter*) or superlative (*heaviest*/*lightest*) form. Similarly, nouns are only sampled when occurring in the singular, so plural pairs such as *men*/*women* and *strengths*/*weaknesses* are not included in the database. Also, verbs are only considered in their base form, so although *win*/*lose* is sampled, *won*/*lost*, *winning*/*losing* and *wins*/*loses* all remain uninvestigated, as do reciprocal verbs such as *buy*/*sell* and *teach*/*learn*, and antonymous prepositions such as *for*/*against* and *to*/*from*. Furthermore, commentators such as Fellbaum (1995: 289) point out that antonyms often co-occur across word class.[7] As such, the antonymy arising between *succeed*/*failure*, *successes*/*fail*, *succeeding*/*failed*, etc. might be explored. It should also be noted that this research deals exclusively with intra-sentential occurrences of antonymy. But antonyms also co-occur inter-sententially so the question remains: how successfully could the functions of antonymy identified here be used to describe how antonymy operates across whole texts? Finally, as mentioned in the previous section, this study restricts itself to analysing the function of antonymy in written text. A corpus of spoken data would show whether (and in what ways) antonymy is used differently in speech. It would also be of interest to discover how successfully the classes developed here would be able to handle the functions of antonymy in other languages.

But perhaps the most pressing area for future research is that of automatic antonym retrieval. The issue of productivity was addressed in Chapter 10 following the discovery that most functions of antonymy are closely associated with certain lexico-syntactic frameworks. For example, the construction *from X to Y* is exclusively associated with Transitional Antonymy, the construction *X not Y* with Negated Antonymy and the construction *more X than Y* with Comparative Antonymy. The association between framework and function is stronger in some cases than others; almost all Distinguished Antonymy sentences rely on *between X and Y*, but Coordinated Antonymy sentences use a plethora of frameworks such as *both X and Y*, *neither X nor Y* and *X and Y alike*. A trio of frameworks were tested by placing a seed word in X-position and examining Y-position output to discover whether this output was contrastive. In general, it was found that these frameworks were successful in retrieving a range of terms, the majority of which were, indeed, contrastive. This enabled

antonymous profiles to be developed for seed words, namely *natural* (which was found to contrast most frequently with *artificial* and *man-made* in text) and *style* (which contrasted with *substance* and *content*). This is an area of research which is especially amenable to further investigation, with the opportunity for text-based dictionaries of antonymy to be developed. From a wider semantic (or even linguistic) perspective, the possibility of being able automatically to identify the antonymous profile of a given word is an exciting step forward.

So what is antonymy then?

The problem of finding an accurate and workable definition of antonymy has been lurking throughout this book. However, a new definition can now be recommended, though this definition must pick up a number of threads. For example, it must account for the fact that antonymy is both a semantic and a lexical relation. Antonymy is lexical because only some word pairs on a given scale are identifiable as 'opposites' (on the scale of height, this lexis is *tall* and *short*, not *lofty* and *petite*). This is partly explained by Carter's notion of 'core vocabulary' (1987) and partly explained by Muehleisen's notion of 'shared semantic range' (1997: iii). Antonymy is semantic for more obvious reasons – word pairs need to belong to the same scale and occupy opposite halves of that scale.

A new definition of antonymy should also incorporate function as a criterion. Table 4.1 shows that all fifty-six antonymous pairs function in a coordinated fashion at least once in the database, and that all but two function in an ancillary fashion at least once in the database. On average, antonyms function in a coordinated fashion at a rate of 38.4 per cent and in an ancillary fashion at a rate of 38.7 per cent. As these distributions are relatively consistent across different pairs, one would not expect to find many antonymous pairs which fail to serve these functions in text.

Finally, a new definition should acknowledge that the litmus test for antonyms is that they be intuitively available to native speakers. Antonymy differs from other sense relations in that it is readily exemplified because of our familiarity with 'opposites'. Any definition of antonymy which is not compatible with the consensus of native-speaker intuition is, in the opinion of this analyst, inherently flawed.

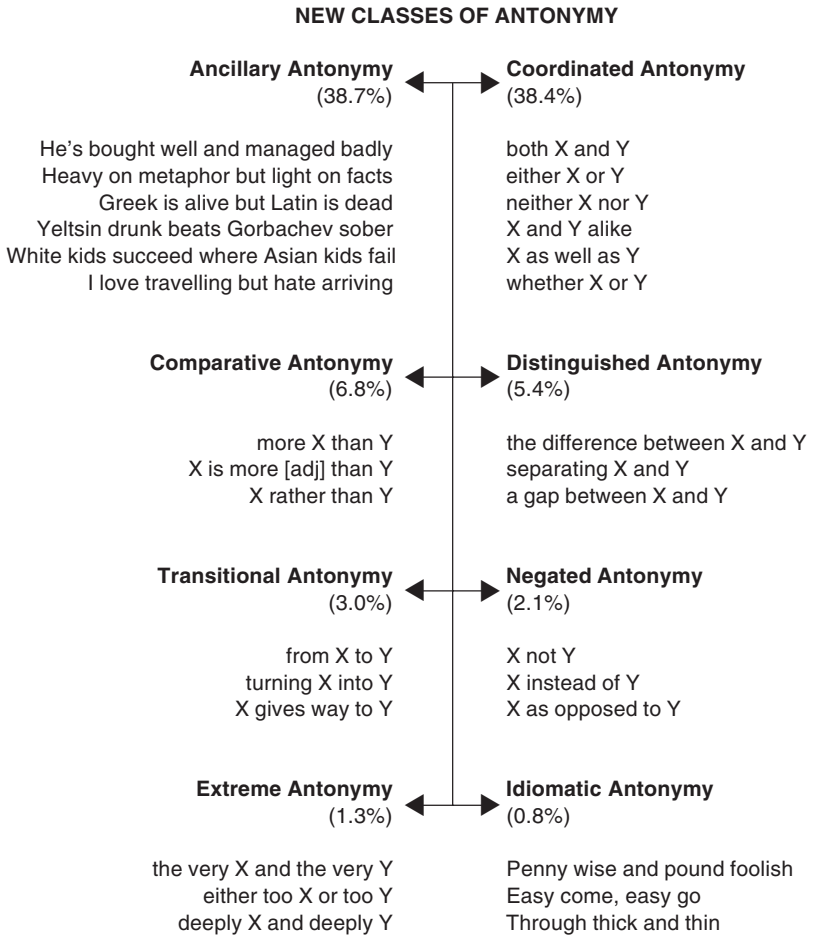
Among criteria which need not be included in a new definition is anything relating to word class, gradability or morphology. The function of antonymy in text remains similar, regardless of whether those pairs are adjectives, nouns, verbs or adverbs; function is also largely unaffected by whether antonyms are gradable or ungradable; and whether antonyms are morphological or lexical. Criteria of this kind are unnecessary because antonymous pairs serve similar functions in text irrespective of their grammatical, logical and morphological composition. Moreover, criteria of this kind are unnecessary because it is possible to identify word pairs of almost any permutation which would be immediately recognised as 'opposites' by native speakers: e.g. *alive/dead*

(adjectival, non-gradable, lexical); *directly/indirectly* (adverbial, morphological), etc.

So, the parting shot of this book is to present a new definition of antonymy which, it is hoped, may prove useful to future antonymists:[8]

> Antonyms are pairs of words which contrast along a given semantic scale and frequently function in a coordinated and ancillary fashion such that they become lexically enshrined as 'opposites'.

Appendix


NEW CLASSES OF ANTONYMY
Ancillary Antonymy
(38.7%)
He's bought well and managed badly
Heavy on metaphor but light on facts
Greek is alive but Latin is dead
Yeltsin drunk beats Gorbachev sober
White kids succeed where Asian kids fail
I love travelling but hate arriving
Coordinated Antonymy
(38.4%)
both X and Y
either X or Y
neither X nor Y
X and Y alike
X as well as Y
whether X or Y
Comparative Antonymy
(6.8%)
more X than Y
X is more [adj] than Y
X rather than Y
Distinguished Antonymy
(5.4%)
the difference between X and Y
separating X and Y
a gap between X and Y
Transitional Antonymy
(3.0%)
from X to Y
turning X into Y
X gives way to Y
Negated Antonymy
(2.1%)
X not Y
X instead of Y
X as opposed to Y
Extreme Antonymy
(1.3%)
the very X and the very Y
either too X or too Y
deeply X and deeply Y
Idiomatic Antonymy
(0.8%)
Penny wise and pound foolish
Easy come, easy go
Through thick and thin

Notes

1 The 'unique fascination' of antonymy

1 So far today, I've listened to a true-or-false quiz on the radio; been told by an advertising hoarding that the *Lancashire Evening Post* is now 'easier to pick up; harder to put down'; noticed that my local cinema is screening a film called *High Heels, Low Lives*; used computer hardware and software; flicked dozens of on/off switches; and had a conversation about the book *Men Are From Mars; Women Are From Venus*. It is no exaggeration to say that antonyms are a ubiquitous part of everyday language and culture.

2 It would seem that once children become familiar with the concept of semantic contrast, they are able to apply the theory to any word (of which an antonym could reasonably be cited) and deduce potential 'opposites'.

3 Lyons (1977) and Cruse (1986) argue that the former group of pairs are 'gradable' (and therefore antonymous), but that the latter group of pairs are 'non-gradable' (and therefore not antonymous). For more detailed discussion of these terms, see Chapter 2.

4 Deese's methodology is discussed more critically on pp. 27–8.

5 Traffic lights may not express antonymy linguistically, but the red and green lights still symbolise underlying contrastive concepts.

6 Cruse states that a relation of sorts holds between all words, sarcastically coining the term 'dogbanonymy' (2000: 145) to describe the sense relation between *dog* and *banana*. Presumably, *happy* and *cutlery* belong to a similar, unique relation, but not one that is recurrent or interesting. For an overview of what makes a sense relation significant, see Cruse (2000: 145–7).

2 A brief history of antonymy

1 For a discussion of the etymology of antonymy, see Muehleisen (1997) or 'Introductory Matter' in *Webster's Dictionary of Synonyms* (1951: vii–xxxiii).

2 Crystal adds the disclaimer: 'the use of the term antonymy must always be viewed with caution' (1985: 18).

3 Most of the research into this quartet of words (*large*, *small*, *big* and *little*) has been undertaken by American linguists, who seem confident that native speaker intuition is sufficient to pair *little* with *big*, and *small* with *large*. However, my instinct is that the antonymous relations holding between these words are less fixed. Could it be that, unlike American researchers, my mental lexicon has been contaminated by childhood exposure to British TV comedy duo Little and Large?

4 Muehleisen tends to avoid the term 'collocation', but notes that *large*/*small* are prone to describe abstract quantities such as *amount* and *area*, and that *big* is often used to describe *news* or *problems* (1997: 68–115).

5 This definition of *passive* also raises the question of how componential analysis would account for *inactive*.
6 Full details of the corpus are provided on p. 26.
7 See pp. 67–8 for a corpus-based reappraisal of the unmarked antonym.
8 Gairns and Redman (1986: 26) report that *cousin* is gender-marked in both Spanish and Portuguese, except when pluralised.
9 The increasing use of *partner* to refer to one's boyfriend, girlfriend, husband or wife is a possible counter-example to the trend of symmetrical gender relations in English being expressed exclusively by formal vocabulary.
10 Conversely, one could easily imagine a hospital patient for whom no doctor has yet been assigned.
11 See Levinson (1983: 140–6) or Hurford and Heasley (1983: 191–7) for more details about the connection between antonymy and truth conditions.
12 Sentences retrieved from the Brown corpus and cited by Justeson and Katz (1991) are analysed on pp. 172–3.
13 This circularity arises because Justeson and Katz are dangerously close to arguing that antonyms are words which are used in antonymous frameworks, and antonymous frameworks are frameworks in which antonyms are used.

3 Approaching antonymy afresh: issues of data and methodology

1 The corpus is genre-specific in that text from only one newspaper source is used, but not in the sense that the newspaper is restricted in its coverage like, say, the *Financial Times*.
2 Deese's list includes words such as *above*, *inside* and *bottom* which function as adjectives less often than they function as other parts of speech.
3 In the case of *gay* and *sad*, it is difficult to imagine that any contrastive context could be contemporary, as both words have shifted in meaning so significantly of late.
4 Reciprocal verbs are not included in the database. Though their textual profile would be interesting to explore, they differ fundamentally from other pairs, in that negating one antonym does not create semantic parity with the other – to *not buy* a product is not the same as to *sell* it.
5 I concur entirely with Fellbaum, who notes that 'there is nothing special about antonymous adjectives, other than that antonymy is more pervasive among adjectives; rather, there is something special about antonymous concepts, no matter in what form these concepts are lexicalised' (1995: 285).
6 These quotas are self-determined but approximate to the distribution of antonyms across language. My own experiments indicate that, when asked to list as many 'opposites' as possible, most informants will supply adjectival pairs at a much higher rate than pairs belonging to other word classes.
7 The instruments used to extract sentences from this corpus were designed by colleagues at the Research and Development Unit for English Studies of Liverpool University. These tools enabled entire sentences from the corpus to be retrieved according to specified search words which co-occurred therein. In this way, all corpus sentences which feature both members of any given antonymous pair could be extracted.
8 The 156 database sentences which feature *un*-words were retrieved from the corpus by software which identified any context in which a word co-occurred with the *un*-version of itself. These sentences were then sampled in accordance with the same procedure used for other antonymous pairs.
9 Our faith in the antonymous prefix is exemplified by the synonyms *flammable* and *inflammable*, which are often regarded as 'opposites' because of their misleading morphology.

4 New classes of antonymy I: Ancillary Antonymy

1 This label should not imply that B-pairs are secondary in any sense; 'B-pair' is not analogous with 'B-movie' or 'B-team'.
2 Synonyms are intuitively regarded as being diametrically opposed to antonyms, but overlaps occasionally arise. For example, within the field of accountancy, (tax) *avoidance* and *evasion* are effectively 'opposites'. This suggests that, at a local level, antonymy can be expressed by terms which would not be interpreted contrastively elsewhere, and may even be regarded as synonyms in other contexts.
3 The term 'meronymy' is here used as a generalisation to describe three slightly different relationships: a *chapter* always has an *end*; a *brothel* must sometimes have *whores*; but a *pond* need never have *fish*.
4 At the point of writing, it is safe to say that the B-pairs identified in sentences 14a–c do not have as high a degree of antonymity as their corresponding A-pairs. However, this need not always be the case, for one of the ways in which antonyms become enshrined in language is through repeated exposure in contexts such as these.
5 Salkie (1995: 77) discusses the role of adversative conjunctions in more detail, though he prefers to describe them as opposition connectives.
6 Sentence 20b is interesting because its B-pair terms (*many* and *a significant number*) would be interchangeable in most contexts, yet are here set up in opposition.
7 'Given' and 'new' can here be simply defined as 'not assumed known' and 'assumed known' respectively (Winter 1982: 110).

5 New classes of antonymy II: Coordinated Antonymy

1 The term 'coordinated' refers to all sentences assigned to the class of Coordinated Antonymy, including those which occur in more grammatically complex constructions such as *X as well as Y*.
2 A normal distribution score of 0.5 indicates no strong tendency to either favour or disfavour the given class; the higher the score, the more marked the inclination for antonyms to serve that particular textual function.
3 The issue of whether word class affects the function of antonymy is discussed further in Chapter 9, but Table 5.1 suggests that adjectives may not favour Coordinated Antonymy as strongly as certain non-adjectival pairs.
4 Fellbaum is describing the encompassing qualities of antonyms when she notes that 'referring to the salient antonymic values of [an] attribute can have the effect of denoting the entire range of values, even though the antonyms may not be the endpoints of the scale' (1995: 295).
5 'Speakers may use both antonyms to eliminate any implications that might arise from the use of the marked or the unmarked term only' (Fellbaum 1995: 299).
6 Extreme Antonymy is analysed on pp. 91–3.
7 *Good* is the 'surprise antonym' in this sentence because one would expect *the bad* to be 'purged', but not *the good*.
8 The parallelism of sentence 35b (*we may succeed, we may fail*) is reminiscent of that identified in Ancillary Antonymy contexts (e.g. sentence 5e: *form is temporary, class is permanent*), but no B-pair contrast is generated.
9 The obvious antonym of *present* would be *absent*, but it is possible that *present* is used to mean 'living' in sentence 36a, thereby rendering that antonym unsuitable.

6 New classes of antonymy III: minor classes

1 This interpretation seems to be favoured by Fellbaum, who handles the framework *X rather than Y* alongside other examples of 'redundancy' such as *X not Y* (1995: 296).

2 The notion of the 'surprise antonym' is introduced on p. 65 and relates to phrases such as *priests, male and female*.
3 The two database sentences which refer to Kipling's quote have actually been classified in terms of Idiomatic Antonymy.
4 Sentences belonging to the class of Distinguished Antonymy have a two-thirds chance of being metalinguistic and a one-third chance of being metaphoric.
5 Political rhetoric is discussed by Heavens (1993) and Fairclough (2000), among others.
6 Strictly speaking, the antonymous pair of sentence 54c is not part of Mr Renton's quote, but one suspects that he is still being paraphrased when *success* and *failure* are mentioned.
7 The term 'redundancy' is used by Fellbaum to describe some of the frameworks associated with Negated Antonymy (1995: 296). However, this term is inappropriate: the negated antonym may be grammatically dispensable and semantically tautologous, but, rhetorically, it is anything but redundant.
8 Sentence 61c is a strange example: Hsu Chu-chuan is said to be *not completely afraid, and not completely unafraid*. Such a distinction (Hsu Chu-chuan is presumably slightly afraid) may seem alien to English speakers, perhaps because the notion of unafraidness is rare or perhaps because of cultural differences in expression.
9 To talk about this 'function' of antonymy may be misleading. Those seventeen database sentences which use an oblique stroke to divide their antonymous pair reflect a variety of textual functions. Unlike other classes, these sentences are grouped together because of a formal feature, namely their punctuation.
10 Quantitative Ancillary Antonymy is discussed on pp. 50–1 and relates to contexts such as *40 per cent agree, but 45 per cent disagree*.
11 It is no coincidence that sentences 70a–c all feature the same antonymous pair – *male* and *female* are the only word pairs to serve the residual function of specification, probably because they are human attributes and, unlike most adjectival antonyms, receptive to quantification.
12 *Life* and *death* are recognisable as 'opposites' to most native speakers of English, but I wonder whether these words are antonymous at all. Surely *death* is, in fact, a part of *life*, just as *dying* is a part of *living*. This would make the relationship between these terms one of meronymy rather than antonymy.

7 The endemicity of antonymy

1 The approach outlined here assumes a corpus-based methodology. Of course, it would be possible to simply read through a few thousand sentences and make a note of any which feature an antonymous pair. However, I have conducted such experiments from time to time and have reservations. First, it is very difficult to read in such an artificial way: if the chosen texts are too interesting, one forgets the purpose of the exercise; if the chosen texts are too dull, one loses concentration. Either way, antonymous pairs (hardly a rare or marked feature of language to begin with) often pass unnoticed. Then there is the problem of finding texts small enough to analyse manually, yet large enough to be representative. For these reasons, I favour a corpus approach, even though this too is far from ideal.
2 This formula is also used by Fellbaum (1995: 287) when calculating co-occurrence probability.
3 The cut-off point of five sentences is arbitrary, though antonyms with a raw frequency so low that co-occurrence is expected in fewer than five sentences per 12 million will struggle to generate reliable O/E scores.
4 A glaring exception to this rule is *female*/*male*, a high-frequency pair which are extremely faithful to one another in text. Indeed, this pair succeed in achieving a higher O/E score than twenty-two of the twenty-eight lower-frequency pairs in the database.

5 Pairs with uneven individual frequencies will generate a lower score for W2/O than for W1/O; words with more even individual frequencies will generate a similar score for both W2/O and W1/O. However, as most antonymous pairs do not have similar individual frequencies, a W1/O score will secure a higher 'league position' than a numerically similar W2/O score. Hence, *rural*/*urban*'s W2/O score of 15.4 fails to make the top ten (Table 7.5), but its W1/O score of 16.7, though numerically similar, is enough to justify third place (Table 7.6).
6 Cruse identifies interesting semantic criteria for 'good opposites' (1986: 262), but co-occurrence data is not used.

8 Antonym sequence

1 For the purpose of this exercise, 'marked preference' is defined as a bias significant enough to generate a binomial score of 0.95 or over.
2 The normal sequence of *quickly*/*slowly* is observed at a statistically significant rate, but the normal sequence of *fast*/*slow* is not.
3 Antonymous pairs cannot always be neatly divided into positive member and negative member – pairs such as *female*/*male*, *cold*/*hot* and *private*/*public* seem largely unbiased.
4 Being *drunk* may sometimes be preferable to being *sober*, but one would expect it carry negative connotations of loutishness and incompetence in British newspaper text.
5 Only five corpus sentences were suitable for this experiment: two showed *selfish* preceding *unselfish*, one showed *unselfish* preceding *selfish*, and two showed *contaminated* preceding *uncontaminated*. In other words, four out of five examples placed their root word first, in spite of its negative associations.
6 It should be noted that *male* and *female* may not be morphologically related at all – this could be an example of folk etymology, with *male* and *female* being derived from the distinct Latin terms *masculus* and *femina*.
7 Lyons applies the theory of 'positive antonym; negative antonym' very liberally, arguing that a positive term can be identified in irreversible binomials such as *men and women* and *food and drink* (1977: 276).
8 A counter-example to gender sequencing is provided by *Mum and Dad*, an antonymous pair which tend to follow a female–male sequence.
9 The thirty-six word pairs which tend to place their higher-frequency antonym first include all nine morphologically derived pairs. If these pairs are excluded on the grounds that their sequence is primarily dictated by their morphological composition, the ratio of those pairs which tend to first-mention their higher-frequency antonym against those pairs which tend to first-mention their lower-frequency antonym is an even more marginal 27:20.
10 Cruse notes that markedness can be defined in different ways (2000: 172–3); here it refers to semantic neutrality (*good* is unmarked because 'did you have a good time?' is less biased (and more common) than 'did you have a bad time?').
11 If markedness can be regarded as a relevant factor, then a possible explanation for the sequence of *hot*/*cold* is that *hot* is the unmarked term and *cold* is marked.
12 If frequency can be regarded as a relevant factor, then a possible explanation for the sequence of *public{ly}*/*private{ly}* is that *public* outnumbers *private* at a ratio of 2:1 and *publicly* outnumbers *privately* at a ratio of 5:4.
13 It is surprising to find that *rightly* precedes *wrongly* in 100 per cent of corpus sentences. Reverse sequence sentences must arise occasionally and one wonders how large a corpus would need to be before they begin to crop up.

9 Antonymy, word class and gradability

1 When *rich* and *poor* function as adjectival noun heads (i.e. *the gap between rich and poor*), they are classed as nouns.

2 Fellbaum notes that 'semantically opposed nouns and verbs do not co-occur in parallel constructions as regularly as the antonymous adjectival pairs observed by Justeson and Katz' (1995: 295). This larger-scale research indicates that Fellbaum is right about nouns (normal distribution score for Coordinated Antonymy: 0.12), but wrong about verbs (normal distribution score for Coordinated Antonymy: 0.68).
3 For the benefit of simplicity and comparability, four artificial examples are used to illustrate this point. However, 'real' contexts in which antonyms serve a similar function could be found in any corpus of modern English.
4 Not all of these thirty-one pairs function exclusively as adjectives in the database. However, the statistics presented in this section disregard those sentences in which adjectival antonyms act as noun heads.
5 When teaching antonymy, I have found that a number of students have argued persuasively that certain, bog-standard examples of gradable antonymy should, in fact, be regarded as non-gradable. For example, are *wet* and *dry* really a gradable pair? If something is completely free of water, it is *dry*; if it is not completely free of water, then it is *wet*, to a greater or lesser degree. Any in-between term, such as *moist* or *damp*, actually refers to the degree of wetness. Ambiguity of this nature further illustrates the problematic nature of intuitive classification.
6 Fellbaum notes that 'association data show that these differences [gradable vs. non-gradable] do not seem to play a role in people's mental organization [of antonyms]' (1995: 285). In this instance, the results of elicitation tests seem to be compatible with corpus-based conclusions.

10 Tomorrow's antonyms

1 The class of Distinguished Antonymy is only a fraction the size of Coordinated Antonymy, which is one of the reasons why so many productive frameworks are associated with the latter. Frequency information about productive frameworks of Coordinated Antonymy can be found on pp. 73–4.
2 These frameworks were selected because they commonly place antonyms in X and Y-position and can therefore be seen as reliable signals of opposition. However, they are not necessarily the most productive and it is quite possible that contrastive Y-position output could be better generated by a variety of frameworks, including or excluding this trio.
3 Table 9.2 shows that nouns favour Distinguished Antonymy (recording a normal distribution score of 0.91) more than adjectives (0.55), but that adjectives favour Coordinated Antonymy (0.36) more than nouns (0.12).
4 Another word which serves a similar role in coordinated contexts is *not*, as in children's hide-and-seek cry of 'coming, ready or not!'
5 The *Collins Cobuild Dictionary* (1987) made use of a corpus which was smaller than my own (about 20 million words), but which was not newspaper-specific.
6 The favoured context of *style* and *substance* is Comparative Antonymy ('style over substance').

11 Antonymy: past, present and future

1 Some words, such as *happy*, may have two potential 'opposites' (*sad* and *unhappy*), but even *happy* has more synonyms than antonyms (*delighted*, *ecstatic*, *cheerful*, etc.).
2 See Cruse (1986: 197) for a list of 'non-learned' terms for antonymy in languages other than English.
3 Palmer (1976: 78–85) is a good example of an adjective-centric approach to antonymy.
4 Justeson and Katz (1991) cite twenty-one corpus sentences, excluding those which feature examples of 'relational adjectives' or 'colour adjectives'.

5 This inclusive use of *but* is unusual but not unprecedented. Sentence 36b states: *they don't encourage it, but they don't discourage it either.*

6 More detailed discussion of these ambiguous attributions can be found in earlier chapters: *neither X nor Y* is analysed on pp. 70–1, *rather X than Y* on pp. 78–9, and the sub-class of Equal Comparison on pp. 79–80.

7 Fellbaum notes that 'the adjective *wet* co-occurs as often with the verb *dry* as with its antonymous adjective *dry*' (1995: 291) and this seems a very persuasive argument for more research into antonym co-occurrence across word class. However, closer inspection reveals that the Brown corpus contains only two examples of *dry* (adjective) co-occurring with *wet* (adjective), and two examples of *dry* (adjective) co-occurring with *wet* (verb). These figures clearly need to be recalculated using a larger corpus before conclusions can be drawn about this issue.

8 Ullman refers to students of synonymy as 'synonymists' (1972: 145); the term 'antonymist' is coined by analogy.

Bibliography

Aitchison, J. (1987) *Words In The Mind: An Introduction to the Mental Lexicon*, Oxford: Blackwell.

Allan, K. (1986) *Linguistic Meaning*, London: Routledge & Kegan Paul.

Baldinger, K. (1980) *Semantic Theory*, Oxford: Blackwell.

Bartsch, R. and Vennemann, T. (1972) *Semantic Structures*, Frankfurt: Athenaum Verlag.

Bauer, L. (1983) *English Word Formation*, Cambridge: Cambridge University Press.

Biber, D., Conrad S. and Reppen, R. (1998) *Corpus Linguistics: Investigating Language Structure and Use*, Cambridge: Cambridge University Press.

Birdsong, D. (1995) 'Iconicity, Markedness, and Processing Constraints in Frozen Locutions', in M. Landsberg (ed.) *Syntactic Iconicity and Linguistic Freezes: The Human Dimension*, Berlin: Mouton de Gruyter.

Butler, C. (1985) *Statistics in Linguistics*, Oxford: Blackwell.

Carnap, R. (1946) *Introduction to Semantics*, Cambridge, Mass.: Harvard University Press.

Carter, R. (1987) *Vocabulary*, London: Allen & Unwin.

Carter, R. and McCarthy, M. (1988) *Vocabulary and Language Teaching*, Harlow: Longman.

Chambers Dictionary of Synonyms and Antonyms (1989), Cambridge: Chambers.

Chomsky, N. (1965) *Aspects of the Theory of Syntax*, Cambridge, Mass.: MIT Press.

Cixous, H. (1989) 'Sorties: Out and Out: Attacks/Ways Out/Forays', in C. Besley and J. Moore (ed.) *The Feminist Reader*, London: Macmillan.

Clark, H. H. (1970) 'Word Associations and Linguistic Theory', in J. Lyons (ed.) *New Horizons in Linguistics*, London: Penguin.

Collier, A., Pacey, M. and Renouf, A. (1998) 'Refining the Automatic Identification of Conceptual Relations in Large-scale Corpora', in E. Charniak (ed.) *Proceedings of the Sixth Workshop on Very Large Corpora*, University of Montreal, Montreal.

Collins Cobuild English Dictionary (1987), London: HarperCollins.

Cook, G. (1989) *Discourse*, Oxford: Oxford University Press.

Cruse, D. A. (1986) *Lexical Semantics*, Cambridge: Cambridge University Press.

——(2000) *Meaning in Language*, Oxford: Oxford University Press.

Crystal, D. (1985) *A Dictionary of Linguistics and Phonetics*, Oxford: Blackwell.

Deese, J. (1964) 'The Associative Structure of Some Common English Adjectives', *Journal of Verbal Learning and Verbal Behaviour*, 3: 347–57.

Dik, S. C. (1980) *Studies in Functional Grammar*, London: Academic Press.

Fairclough, N. (2000) *New Labour, New Language*, London: Routledge.
Fellbaum, C. (1995) 'Co-occurrence and Antonymy', *International Journal of Lexicography*, 8.4: 281–303.
Fodor, U. D. (1982) *Semantics*, Brighton: Harvester Press.
Frawley, W. (1992) *Linguistic Semantics*, New Jersey: Erlbaum.
Gairns, R. and Redman, S. (1986) *Working with Words*, Cambridge: Cambridge University Press.
Grefenstette, G. (1992) 'Finding Semantic Similarity in Raw Text: the Deese Antonyms', in *AAAI Fall Symposium Series: Probabilistic Approaches to Natural Language*, Cambridge, Mass.
Halliday, M. A. K. and Hasan, R. (1976) *Cohesion in English*, Harlow: Longman.
Hasan, R. (1984) 'Coherence and Cohesive Harmony', in J. Flood (ed.) *Understanding Reading Comprehension*, Newark: International Reading Association.
Hearst, M. A. (1998) 'Automated Discovery of WordNet Relations', in C. Fellbaum (ed.) *WordNet*, Cambridge, Mass.: MIT Press.
Heavens, S. (1993) 'Replacement Grammar in Political Rhetoric', unpublished M.Phil. dissertation, University of Birmingham.
Hoey, M. (1991) *Patterns of Lexis in Text*, Oxford: Oxford University Press.
Hoffman, T. R. (1993) *Realms of Meaning*, Harlow: Longman.
Hurford, J. R. and Heasley, B. (1983) *Semantics: A Coursebook*, Cambridge: Cambridge University Press.
Jackendoff, R. S. (1983) *Semantics and Cognition*, Cambridge, Mass.: MIT Press.
——(1990) *Semantic Structures*, Cambridge, Mass.: MIT Press.
Jackson, H. (1988) *Words and their Meaning*, London: Longman.
Jung, C. G. (1973) *Collected Works of C. G. Jung*, vol. 2, Princeton: Princeton University Press.
Justeson, J. S. and Katz, S. M. (1991) 'Redefining Antonymy: The Textual Structure of a Semantic Relation', *Literary and Linguistic Computing*, 7: 176–84.
Kagan, J. (1984) *The Nature of the Child*, New York: Basic Books.
Katamba, F. (1994) *English Words*, London: Routledge.
Katz, J. J. (1972) *Semantic Theory*, New York: Harper & Row.
Kempson, R. U. (1977) *Semantic Theory*, Cambridge: Cambridge University Press.
Kennedy, G. (1998) *An Introduction to Corpus Linguistics*, Harlow: Longman.
Kimball, J. P. (1975) *Syntax and Semantics*, New York: Academic Press.
Kwon, H.-S. (1998) 'A Corpus-based Study of English Negative Prefixation', in *1998 TALC Proceedings*, London: Longman.
Lakoff, G. and Johnson, M. (1980) *Metaphors We Live By*, Chicago: University of Chicago Press.
Leech, G. (1974) *Semantics*, Harmondsworth: Penguin.
Lehrer, A. (1974) *Semantic Fields and Lexical Structure*, Amsterdam: North Holland Publishing Company.
Levinson, S. C. (1983) *Pragmatics*, Cambridge: Cambridge University Press.
Lyons, J. (1977) *Semantics*, 2 vols., Cambridge: Cambridge University Press.
——(1981) *Language, Meaning and Context*, London: Fontana Paperbacks.
McCarthy, M. (1990) *Vocabulary*, Oxford: Oxford University Press.
——(1991) *Discourse Analysis for Language Teachers*, Cambridge: Cambridge University Press.
——(1998) *Spoken Language and Applied Linguistics*, Cambridge: Cambridge University Press.

McEnery, T. and Wilson, A. (1996; 2nd edn 2001) *Corpus Linguistics*, Edinburgh: Edinburgh University Press.

Mettinger, A. (1994) *Aspects of Semantic Opposition in English*, Oxford: Oxford University Press.

Muehleisen, V. (1997) 'Antonymy and Semantic Range in English', unpublished Ph.D. dissertation, Northwestern University.

Murphy, L. M. (1994) 'In Opposition to an Organised Lexicon: Pragmatic Principles and Lexical Semantic Relations', unpublished Ph.D. dissertation, University of Illinois.

Nesfield, J. C. (1898) *Manual of English Grammar and Composition*, Toronto: Macmillan.

Oakes, M. P. (1998) *Statistics for Corpus Linguistics*, Edinburgh: Edinburgh University Press.

Palmer, F. R. (1976) *Semantics*, Cambridge: Cambridge University Press.

Quirk, R., Greenbaum, S., Leech, G. and Svartvik, J. (1972) *A Grammar of Contemporary English*, London: Longman.

Renouf, A. (1996) 'The ACRONYM Project: Discovering the Textual Thesaurus', in J. M. Aarts, P. De Haan and N. H. J. Oostdijk (eds) *Synchronic Corpus Linguistics – Papers from the Sixteenth International Conference on English Language Research on Computerised Corpora*, Amsterdam: RoDoPi.

Renouf, A. and Baayen, H. (1996) 'Chronicling the Times: Productive Lexical Innovations in an English Newspaper', *Language*, 72.1: 69–96.

——(1998) 'Aviating Among the Hapax Legomena', in A. Renouf (ed.) *Explorations in Corpus Linguistics*, Amsterdam: RoDoPi.

Richard, J. C., Platt, J. and Webber, H. (1985) *Longman Dictionary of Applied Linguistics*, Harlow: Longman.

Roget's Thesaurus (1952) London: Sphere Books.

Rosenthal, P. (1984) *Words and Values*, Oxford: Oxford University Press.

Rudska, B., Channell, J., Putseys, Y. and Ostyn, J. (1981) *The Words You Need*, London: Macmillan.

Salkie, R. (1995) *Text and Discourse Analysis*, London: Routledge.

Simpson, P. (1997) *Literature Through Language*, London: Routledge.

Sinclair, J. McH. (1991) *Corpus, Concordance, Collocation*, Oxford: Oxford University Press.

Stubbs, M. (1996) *Text and Corpus Analysis: Computer Assisted Studies of Language and Culture*, Oxford: Blackwell.

Taylor, M. C. (1984) *Erring: A Postmodern A/Theology*, Chicago: University of Chicago Press.

Ullman, S. (1951) *The Principles of Semantics*, Glasgow: Jackson.

——(1972) *Semantics*, Oxford: Blackwell.

Webster's Dictionary of Synonyms (1951), Measha: Merriam.

Winter, E. (1982) *Towards a Contextual Grammar of English*, London: Allen & Unwin.

Index